

From: Myers, Carla
Sent: Wednesday, July 13, 2005 7:06 AM
To: STIC-Biotech/ChemLib
Subject: sequence search for 08/974,584

(STIC)

JUL 13 2005

RECEIVED

Please search commercial and interference databases for:

- 1) nucleic acids encoding SEQ ID NO: 118 (1132 amino acids);
- 2) proteins comprising SEQ ID NO: 118 (1132 amino acids)
- 3) nucleic acids encoding each of SEQ ID NO: 139, 143, 144, 146, 147, and 16
- 4) nucleic acids encoding each of SEQ ID NO: 139, 143, 144, 146, 147 and 17

** if it is not possible to search 3 and 4 (i.e., nucleic acids encoding multiple SEQ ID NOs), then please perform a separate search for nucleic acids encoding SEQ ID NO: 139, 143, 144, 146, 147, 16 and 17.

SEQ 139=11 amino acids; SEQ 143=5 amino acids; SEQ 144=22 amino acids; SEQ 147=13 amino acids;
SEQ 16=50 amino acids; SEQ 17=43 amino acids

The CRF has been entered: http://expoweb1:8001/cgi-bin/expo/BioInfo/bioquery.pl?APPL_ID=08974584

**Please provide a printout of the first 40 results.

Thank you-

Carla Myers
AU 1634
Remsen Bldg / Rm 2E79
Mailbox: REM 2C70
571-272-0747

61/9 03p/17 03h/13 04/36 05p/4 05h/20
06p/10 06h/11

* myers 08974584*
* us08974584*

STAFF USE ONLY

Searcher: MBB
Searcher Phone: 2-
Date Searcher Picked up: 7-27-05
Date Completed: 7-27-05
Searcher Prep/Rev. Time: 20
Online Time: 18

Type of Search

NA#: 3 AA#: 1
Interference: SPDI:
S/L: Oligomer:
Encode/Trans: Text:
Structure#: Litigation:
Inventor: Litigation:

Vendors and cost where applicable

STN: 06h, IG
DIALOG: IG
QUESTEL/ORBIT: IG
LEXIS/NEXIS: IG
SEQUENCE SYSTEM: IG
WWW/Internet: IG
Other(Specify): IG

THIS PAGE BLANK (USPTO,

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 25, 2005, 22:18:34 ; Search time 4179.29 Seconds

(without alignments)
938.107 Million cell updates/sec

Title: SEQ139-143-144-146-147-17

Perfect score: 233

Sequence: 1 XXXRXXPKXXXRXIXXXXX.....WXTXXXXXXXRXRXXXW 103

Scoring table:

BLSUM62
Xgapop 10.0 , Xgapext 0.1
Ygapop 10.0 , Ygapext 0.1
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPTO spo01/MYERS08974584/runat_25072005_102706_5075/app_query.fasta_1.526
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=MYERS08974584 @CGN 1.1 3556 @runat_25072005_102706_5075 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT_DSPLCK=100 -JONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THRADS=1 -XGAPOP=10 -XGAPEXT=0.1 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.1 -DELOP=6 -DELEXT=7

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hcc:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query ID	Description
C 1	39	16.7	352	BP427888
C 2	39	16.7	437	BUS06243
C 3	39	16.7	576	BP760208
C 4	39	16.7	601	BP766928
C 5	39	16.7	642	BP771659
C 6	38	16.3	1047	BQ224762
C 7	37	15.9	939	CL398497
C 8	36.8	15.8	851	BUS09508
C 9	36.4	15.6	1058	BH725011
C 10	36.1	15.5	181	CG784907
C 11	36.1	15.5	197	CG887155
C 12	36.1	15.5	216	CB178950
C 13	36.1	15.5	357	CG784906
C 14	36.1	15.5	408	CK225321
C 15	36.1	15.5	563	CK225321
C 16	36.1	15.5	578	CK722398
C 17	36.1	15.5	624	CK722398
C 18	36.1	15.5	665	CK722398
C 19	36.1	15.5	741	CK722398
C 20	36.1	15.5	746	CK780407
C 21	36.1	15.5	748	CK780407
C 22	36.1	15.5	768	CK780407
C 23	36.1	15.5	859	CK780407
C 24	36.1	15.5	939	CK780407
C 25	36.1	15.5	1024	CK780407
C 26	36.1	15.5	1024	CK780407
C 27	36.1	15.5	1024	CK780407
C 28	36.1	15.5	1024	CK780407
C 29	36.1	15.5	1024	CK780407
C 30	36.1	15.5	1024	CK780407
C 31	36.1	15.5	1024	CK780407
C 32	36.1	15.5	1024	CK780407
C 33	36.1	15.5	1024	CK780407
C 34	36.1	15.5	1024	CK780407
C 35	36.1	15.5	1024	CK780407
C 36	36.1	15.5	1024	CK780407
C 37	36.1	15.5	1024	CK780407
C 38	36.1	15.5	1024	CK780407
C 39	36.1	15.5	1024	CK780407
C 40	36.1	15.5	1024	CK780407
C 41	36.1	15.5	1024	CK780407
C 42	36.1	15.5	1024	CK780407
C 43	36.1	15.5	1024	CK780407
C 44	36.1	15.5	1024	CK780407
C 45	36.1	15.5	1024	CK780407
C 46	36.1	15.5	1024	CK780407
C 47	36.1	15.5	1024	CK780407
C 48	36.1	15.5	1024	CK780407
C 49	36.1	15.5	1024	CK780407
C 50	36.1	15.5	1024	CK780407
C 51	36.1	15.5	1024	CK780407
C 52	36.1	15.5	1024	CK780407
C 53	36.1	15.5	1024	CK780407
C 54	36.1	15.5	1024	CK780407
C 55	36.1	15.5	1024	CK780407
C 56	36.1	15.5	1024	CK780407
C 57	36.1	15.5	1024	CK780407
C 58	36.1	15.5	1024	CK780407
C 59	36.1	15.5	1024	CK780407
C 60	36.1	15.5	1024	CK780407
C 61	36.1	15.5	1024	CK780407
C 62	36.1	15.5	1024	CK780407
C 63	36.1	15.5	1024	CK780407
C 64	36.1	15.5	1024	CK780407
C 65	36.1	15.5	1024	CK780407
C 66	36.1	15.5	1024	CK780407
C 67	36.1	15.5	1024	CK780407
C 68	36.1	15.5	1024	CK780407
C 69	36.1	15.5	1024	CK780407
C 70	36.1	15.5	1024	CK780407
C 71	36.1	15.5	1024	CK780407
C 72	36.1	15.5	1024	CK780407
C 73	36.1	15.5	1024	CK780407
C 74	36.1	15.5	1024	CK780407
C 75	36.1	15.5	1024	CK780407
C 76	36.1	15.5	1024	CK780407
C 77	36.1	15.5	1024	CK780407
C 78	36.1	15.5	1024	CK780407
C 79	36.1	15.5	1024	CK780407
C 80	36.1	15.5	1024	CK780407
C 81	36.1	15.5	1024	CK780407
C 82	36.1	15.5	1024	CK780407

83	33	14.2	156	1	AL797816	AL797816
84	33	14.2	165	1	AV157928	AV157928
85	33	14.2	171	1	AL883683	AL883683
86	33	14.2	171	1	AL883683	AL883683
87	33	14.2	173	2	BE089816	RC5-BT070
88	33	14.2	176	2	BF79876	BF79876
89	33	14.2	184	8	AZ338728	RC2-ET018
90	33	14.2	191	2	BF118430	AZ338728
91	33	14.2	197	1	AI609913	1U0669J21
92	33	14.2	211	6	CB189535	BF118430
93	33	14.2	216	1	AL868343	IM95d09.x
94	33	14.2	220	1	AAZ25807	AI609913
95	33	14.2	229	7	CF735736	CF209635
96	33	14.2	233	2	BB132151	CF209635
97	33	14.2	240	2	BE117386	CB189535
98	33	14.2	243	7	CV321951	AL868343
99	33	14.2	246	5	BX629569	AAZ25807
100	33	14.2	246	7	CV323271	CF735736
101	33	14.2	246	7	CV323271	UI-M-HB0
102	33	14.2	246	7	CV323271	BB132151
103	33	14.2	246	7	CV323271	UI-R-BB1
104	33	14.2	246	7	CV323271	BE117386
105	33	14.2	246	7	CV323271	CV321951
106	33	14.2	246	7	CV323271	CM3-HT096
107	33	14.2	246	7	CV323271	BX629569
108	33	14.2	246	7	CV323271	CM4-AN007

ALIGNMENTS

RESULT	1
BP427868/c	
LOCUS	352 bp mRNA linear EST 24-MAY-2004
DEFINITION	Mus musculus cerebellum E18-P56 Mus musculus cDNA clone FP1633, mRNA sequence.
ACCESSION	BP427868
VERSION	BP427868.1
KEYWORDS	GI:47589962
SOURCE	EST.
ORGANISM	Mus musculus (house mouse)
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus. 1 (bases 1 to 352) Furuichi,T. and Sato,A. Gene expression during the postnatal cerebellar development Unpublished (2004) Contact: Teiichi Furuichi Laboratory for Molecular Neurogenesis RIKEN Brain Science Institute 2-1 Hiroawa, Wako, Saitama 351-0198, Japan Tel: 81-48-467-5906 Fax: 81-48-467-6079 Email: mol-neurogenesis@brain.riken.go.jp, URL: http://www.brain.riken.go.jp/labs/lmn/index.html .

903-877-3333

Alignment Scores:		
Pred. No.:	164	Length:
Score:	39.00	Matches:
Percent Similarity:	31.58%	Conservative:
Best Local Similarity:	31.58%	Mismatches:
Query Match:	16.74%	Indels:
DB:	5	Gaps:
		0

SEO139-143-144-146-147-17 (1-103) X BP427888 (1-352)

Qy 69 TTP*****PhePheTrp***ThrGlu 87
 ||| {} ||||| |||||
 Db 103 TGGAGTTCAGATCCATCGCAGCAAGTCTCAGATCTCTTTCTGGGAACAGAA 47

RESULT 2

```

BU506243/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
1
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
    source
        BU506243
        AGENCOURT 10047055 NIH_MGC_94 Mus musculus cDNA clone IMAGE:6494302
        S', mRNA sequence.
        BU506243
        BU506243.1 GI:22812476
        Mus musculus (house mouse)
        Mus musculus
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
        1 (bases 1 to 437)
        NIH-MGC http://img.nci.nih.gov/.
        National Institutes of Health, Mammalian Gene Collection (MGC)
        Unpublished (1999)
        Contact: Robert Strausberg, Ph.D.
        Email: cgabbs@mail.nih.gov
        Tissue Procurement: The Cepko Laboratory
        cDNA Library Preparation: Life Technologies, Inc.
        cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
        DNA Sequencing by: Agencourt Bioscience Corporation
        Clone distribution: MGC clone distribution information can be
        found through the I.M.A.G.E. Consortium/LLNL at:
        http://image.llnl.gov
        Plate: LLAM14050 row: e column: 23
        High quality sequence stop: 436.
        Location/Qualifiers
            1..437
                /organism="Mus musculus"
                /mol_type="mRNA"
                /db_xref="taxon:10090"
                /clone="IMAGE:6494302"
                /tissue_type="retina"
                /lab_host="DH10B (phage-resistant)"
                /clone_lib="NIH MGC 94"
                /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
                Site 2: SalI; Cloned unidirectionally; oligo-dt primed.
                Average insert size 3.3 kb. Library enriched for
                full-length clones and constructed by Life Technologies.
                Note: this is a NIH MGC library."

```

ORIGIN

Alignment Scores:		
Pred. No.:	207	Length: 437
Score:	39.00	Matches: 6
Percent Similarity:	31.58%	Conservative: 0
Best Local Similarity:	31.58%	Mismatches: 138
Query Match:	16.74%	Indels: 0
DB:	5	Gaps: 0

SEO139-143-144-146-147-17 (1-103) x BU506243 (1-437)

Qy 69 Trp*****PhePheTrp***ThrGlu 87
|||
Db 138 TGGAGTTCAGATTCATCGCAGCCAGTCTCCAGTCTCTTTCTGGGAACAGAA 82

BF700205	BP760208	576 bp	mRNA	linear	EST 10-JUL-2004
LOCUS	BP760208	mouse (C57BL/6)	pancreatic islet	library with	
DEFINITION	recombination-based method Mus musculus cDNA clone mib31028 3', mRNA sequence.				

RESULT 2

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 26, 2005, 00:55:25 ; Search time 1188.27 Seconds
(without alignments)
560.502 Million cell updates/sec

Title: SEQ139-143-144-146-147-17
Perfect score: 233
Sequence: 1 XXXRXXPPXXXXXXIXXXXX.....WXTXXXXXXXIXXXXXX 103

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.1
Ygapop 10.0 , Ygapext 0.1
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7277826 seqs, 3233139505 residues

Total number of hits satisfying chosen parameters: 14555652

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/MYERS08974584/runat_25072005_102708_5180/app_query.fasta_1.526
-DB=Published Applications NA -QPMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MAIRIX=BLOSUM62
-TRANS=human40.cdi -LIST=100 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=40 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000
-USER=MYERS08974584 @CGN 1.1.456 @runat_25072005_102708_5180 -NCPUs=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THRAOPS=1 -XGAPOP=10 -XGAPEXT=0.1 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.1 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	35.9	15.4	368	9	US-09-770-791-397	Sequence 397, App
2	35.9	15.4	434	9	US-09-924-035A-656	Sequence 656, App
3	35.9	15.4	639	9	US-09-938-842A-2291	Sequence 2291, App
4	35.9	15.4	639	11	US-09-938-842A-2291	Sequence 2291, App
5	35.9	15.4	2321	20	US-10-739-963-1097	Sequence 1097, App
C 6	35	15.0	455	19	US-10-437-963-9511	Sequence 9511, App
7	35	15.0	600	22	US-10-972-079-62547	Sequence 62547, App
8	34.1	14.6	784	17	US-10-225-066A-655	Sequence 655, App
9	34.1	14.6	784	17	US-10-374-780A-2237	Sequence 2237, App
10	34	14.6	404	18	US-10-424-599-40084	Sequence 40084, App
C 11	33	14.2	102	21	US-10-472-928-1683	Sequence 1683, App
C 12	33	14.2	140	9	US-09-783-590-278	Sequence 278, App
13	33	14.2	193	19	US-10-437-963-102089	Sequence 102089,
14	33	14.2	201	20	US-10-719-993-25917	Sequence 25917, App
C 15	33	14.2	282	18	US-10-424-599-57538	Sequence 57538, App
16	33	14.2	326	19	US-10-767-795-4667	Sequence 4667, App
C 17	33	14.2	330	9	US-09-783-590-11246	Sequence 11246, App
18	33	14.2	376	9	US-09-880-107-1031	Sequence 1031, App
19	33	14.2	460	20	US-10-425-115-80762	Sequence 80762, App
20	33	14.2	488	14	US-10-060-036-1633	Sequence 1633, App
C 21	33	14.2	521	20	US-10-425-115-36983	Sequence 36983, App
22	33	14.2	529	19	US-10-767-701-5515	Sequence 5515, App
C 23	33	14.2	539	18	US-10-424-599-29238	Sequence 29238, App
C 24	33	14.2	572	19	US-10-021-323-12142	Sequence 12142, App
25	33	14.2	593	13	US-10-027-632-271755	Sequence 271755,
26	33	14.2	593	17	US-10-027-632-271755	Sequence 271755,
C 27	33	14.2	595	19	US-10-767-795-4666	Sequence 4666, App
C 28	33	14.2	616	13	US-10-027-632-265805	Sequence 265805,
C 29	33	14.2	616	17	US-10-027-632-265806	Sequence 265806,
C 30	33	14.2	616	17	US-10-027-632-265805	Sequence 265805,
C 31	33	14.2	616	17	US-10-027-632-265806	Sequence 265806,
32	33	14.2	619	13	US-10-027-632-196891	Sequence 196891,
33	33	14.2	619	13	US-10-027-632-196892	Sequence 196892,
34	33	14.2	619	17	US-10-027-632-196893	Sequence 196893,
35	33	14.2	619	17	US-10-027-632-196891	Sequence 196891,
36	33	14.2	619	17	US-10-027-632-196892	Sequence 196892,
37	33	14.2	619	17	US-10-027-632-196893	Sequence 196893,
C 38	33	14.2	649	20	US-10-335-053-194	Sequence 194, App
C 39	33	14.2	652	20	US-10-357-930-52657	Sequence 52657, App
40	33	14.2	664	18	US-10-425-114-16192	Sequence 16192, App
41	33	14.2	667	13	US-10-027-632-287415	Sequence 287415,
42	33	14.2	667	13	US-10-027-632-287416	Sequence 287416,
43	33	14.2	667	17	US-10-027-632-287415	Sequence 287415,
44	33	14.2	667	17	US-10-027-632-287416	Sequence 287416,
45	33	14.2	721	13	US-10-027-632-260266	Sequence 260266,
46	33	14.2	721	17	US-10-027-632-260266	Sequence 260266,
47	33	14.2	889	20	US-10-363-345A-28829	Sequence 28829, App
C 48	33	14.2	889	20	US-10-363-345A-28830	Sequence 28830, App
C 49	33	14.2	889	21	US-10-363-345A-28829	Sequence 28829, App
C 50	33	14.2	889	21	US-10-363-345A-28830	Sequence 28830, App
C 51	33	14.2	900	20	US-10-425-115-160420	Sequence 160420,
52	33	14.2	987	18	US-10-424-599-90405	Sequence 90405, App
C 53	33	14.2	1113	19	US-10-474-776-76	Sequence 76, Appl
C 54	33	14.2	1162	21	US-10-956-157-1455	Sequence 1455, App
C 55	33	14.2	1162	21	US-10-956-157-6690	Sequence 6690, App
56	33	14.2	1171	17	US-10-369-493-36598	Sequence 36598, App
57	33	14.2	1171	20	US-10-363-345A-40425	Sequence 40425, App
C 58	33	14.2	1174	20	US-10-363-345A-40426	Sequence 40426, App
59	33	14.2	1174	21	US-10-363-345A-40425	Sequence 40425, App
C 60	33	14.2	1174	21	US-10-363-345A-40426	Sequence 40426, App
C 61	33	14.2	1201	20	US-10-425-115-89303	Sequence 89303, App
62	33	14.2	1312	18	US-10-425-114-25327	Sequence 25327, App
63	33	14.2	1383	19	US-10-437-963-100022	Sequence 100022,
64	33	14.2	1404	20	US-10-425-115-49114	Sequence 49114, App

```
65 33 14.2 1554 21 US-10-470-048B-537
66 33 14.2 1575 22 US-10-724-972A-3346
67 33 14.2 1631 13 US-10-027-632-258931
68 33 14.2 1631 13 US-10-027-632-258932
69 33 14.2 1631 13 US-10-027-632-258933
70 33 14.2 1631 13 US-10-027-632-258934
71 33 14.2 1631 17 US-10-027-632-258931
72 33 14.2 1631 17 US-10-027-632-258932
73 33 14.2 1631 17 US-10-027-632-258933
74 33 14.2 1631 17 US-10-027-632-258934
75 33 14.2 1988 13 US-10-027-632-97163
76 33 14.2 1988 13 US-10-027-632-97164
77 33 14.2 1988 13 US-10-027-632-97165
78 33 14.2 1988 17 US-10-027-632-97163
79 33 14.2 1988 17 US-10-027-632-97164
80 33 14.2 1988 17 US-10-027-632-97165
81 33 14.2 2418 9 US-09-938-842A-1478
82 33 14.2 2418 11 US-09-938-842A-1478
83 33 14.2 2563 9 US-09-782-051-1
84 33 14.2 2565 21 US-10-876-285-3
85 33 14.2 2760 20 US-10-739-930-157
86 33 14.2 3634 20 US-10-357-930-23184
87 33 14.2 3634 20 US-10-357-930-29050
88 33 14.2 3986 13 US-10-092-925-1
89 33 14.2 4105 17 US-10-062-674-1991
90 33 14.2 4333 20 US-10-357-930-21849
91 33 14.2 4333 20 US-10-357-930-22736
92 33 14.2 4333 20 US-10-357-930-27699
93 33 14.2 4333 20 US-10-357-930-28584
94 33 14.2 5201 19 US-10-717-597-213
95 33 14.2 5384 15 US-10-311-455-181
96 33 14.2 6515 15 US-10-311-455-379
97 33 14.2 18627 8 US-08-961-527-113
98 33 14.2 18627 17 US-10-158-844-113
99 33 14.2 44096 21 US-10-876-285-1
100 33 14.2 88853 13 US-10-087-192-85

ALIGNMENTS

RESULT 1
US-09-770-791-397
; Sequence 397, Application US/09770791
; Patent No. US20020062014A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Gricker, Carlos A.
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurlan, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2029 (PARA-018PRV)
; CURRENT APPLICATION NUMBER: US/09/770,791
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,480
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0

Sequence 537, App
Sequence 3346, Ap
Sequence 258931,
Sequence 258932,
Sequence 258933,
Sequence 258934,
Sequence 258931,
Sequence 258932,
Sequence 258933,
Sequence 258934,
Sequence 97163, A
Sequence 97164, A
Sequence 97165, A
Sequence 97163, A
Sequence 97164, A
Sequence 97165, A
Sequence 1478, Ap
Sequence 1478, Ap
Sequence 1, Appli
Sequence 3, Appli
Sequence 157, App
Sequence 23184, A
Sequence 29050, A
Sequence 1, Appli
Sequence 1991, Ap
Sequence 21849, A
Sequence 22736, A
Sequence 27699, A
Sequence 28584, A
Sequence 213, App
Sequence 181, App
Sequence 379, App
Sequence 113, App
Sequence 113, App
Sequence 1, Appli
Sequence 85, Appli

; SEQ ID NO 397
; LENGTH: 368
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(368)
; OTHER INFORMATION: n = A,T,C or G
US-09-770-791-397
Alignment Scores:
Pred. No.: 183
Score: 35.90
Percent Similarity: 26.92%
Best Local Similarity: 26.92%
Query Match: 15.41%
Indels: 11
Gaps: 1
Length: 368
Matches: 7
Conservative: 0
Mismatch: 8
Indels: 11
Gaps: 1

SEQ139-143-144-146-147-17 (1-103) x US-09-770-791-397 (1-368)
Qy 62 AspAsp***Leu*****Trp***** 81
Db 231 GACGACTCTCTATCTGCCACGTGG-----CGG 257
Qy 82 PhePheTrp***ThrGlu 87
Db 258 CTCTTGTGGACGACGGAG 275

RESULT 2
US-09-924-035A-656
; Sequence 656, Application US/09924035A
; Patent No. US20020142319A1
; GENERAL INFORMATION:
; APPLICANT: Grilach, Jrn
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2011US
; CURRENT APPLICATION NUMBER: US/09/924,035A
; CURRENT FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: US 60/148,784
; PRIOR FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 900
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 656
; LENGTH: 434
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(434)
; OTHER INFORMATION: n = A,T,C or G
US-09-924-035A-656
Alignment Scores:
Pred. No.: 216
Score: 35.90
Percent Similarity: 26.92%
Best Local Similarity: 26.92%
Query Match: 15.41%
Indels: 11
Gaps: 1
Length: 434
Matches: 7
Conservative: 0
Mismatch: 8
Indels: 11
Gaps: 1

SEQ139-143-144-146-147-17 (1-103) x US-09-924-035A-656 (1-434)
Qy 62 AspAsp***Leu*****Trp***** 81
Db 293 GACGACTCTCTATCTGCCACGTGG-----CGG 319
Qy 82 PhePheTrp***ThrGlu 87
Db 320 CTCTTGTGGACGACGGAG 337

RESULT 3
US-09-938-842A-2291
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 25, 2005, 22:19:19 ; Search time 1974.08 Seconds
(without alignments)
85.375 Million cell updates/sec

Title: SEQ139-143-144-146-147-17

Perfect score: 233

Sequence: 1 XXXRXPXXXXXIXXXXXX.....WXTXXXXXXXXXXXXXXW 103

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.1
Ygapop 10.0, Ygapext 0.1
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/MYERS08974584/runat_25072005_102706_5094/app.query.fasta_1.526
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rn1 -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=100 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=MYERS08974584 @Cn 1.1 77 @runat_25072005_102706_5094 -NCFU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRSADS=1 -XGAPOP=10 -XGAPEXT=0.1 -FGAPOP=6
-XGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.1 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*

5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	33	14.2	601	4	US-09-949-016-68055
C 2	33	14.2	601	4	US-09-949-016-72934
C 3	33	14.2	601	4	US-09-949-016-72935
C 4	33	14.2	601	4	US-09-949-016-206055
C 5	33	14.2	601	4	US-09-949-016-206056
C 6	33	14.2	601	4	US-09-949-016-206092
C 7	33	14.2	601	4	US-09-949-016-206093
C 8	33	14.2	1575	3	US-09-134-003C-2620
C 9	33	14.2	2226	2	US-08-896-003-2
C 10	33	14.2	2226	3	US-09-347-878-3
C 11	33	14.2	18627	3	US-08-961-527-113
C 12	33	14.2	104428	4	US-09-949-016-12737

Sequence 13814, A	US-09-949-016-13814	4	14.2	104429	33
Sequence 17561, A	US-09-949-016-17561	4	14.2	113379	33
Sequence 17562, A	US-09-949-016-17562	4	14.2	113379	33
Sequence 17202, A	US-09-949-016-17202	4	14.2	189560	33
Sequence 11869, A	US-09-949-016-11869	4	14.2	194714	33
Sequence 13875, A	US-09-949-016-13875	4	14.2	221545	33
Sequence 15725, A	US-09-949-016-15725	4	14.2	264358	33
Sequence 13747, A	US-09-949-016-13747	4	14.2	264665	33
Sequence 21066, A	US-09-513-999C-21066	4	14.2	264665	33
Sequence 3, Appl1	US-09-161-241-3	3	14.1	458	32.8
Sequence 391, App	US-09-919-039-391	4	14.1	1215	32.8
Sequence 1085, Ap	US-09-976-594-1085	4	14.1	1797	32.8
Sequence 23306, A	US-09-949-016-23306	4	13.8	601	32.2
Sequence 48571, A	US-09-949-016-48571	4	13.8	601	32.2
Sequence 12003, A	US-09-949-016-12003	4	13.8	48119	32.2
Sequence 13177, A	US-09-949-016-13177	4	13.8	48119	32.2
Sequence 23, Appl1	US-09-018-584A-23	3	13.7	359	32
Sequence 23, Appl1	US-09-784-423-23	3	13.7	359	32
Sequence 55, Appl1	US-08-313-075A-55	1	13.7	386	32
Sequence 13, Appl1	US-09-976-594-13	4	13.7	460	32
Sequence 11994, A	US-09-270-767-11994	4	13.7	462	32
Sequence 444, App	US-09-107-532A-206	4	13.7	462	32
Sequence 10088, A	US-10-101-464A-444	4	13.7	516	32
Sequence 25892, A	US-09-949-016-25892	4	13.7	601	32
Sequence 28910, A	US-09-949-016-28910	4	13.7	601	32
Sequence 31502, A	US-09-949-016-31502	4	13.7	601	32
Sequence 31503, A	US-09-949-016-31503	4	13.7	601	32
Sequence 32526, A	US-09-949-016-32526	4	13.7	601	32
Sequence 46140, A	US-09-949-016-46140	4	13.7	601	32
Sequence 46141, A	US-09-949-016-46141	4	13.7	601	32
Sequence 61322, A	US-09-949-016-61322	4	13.7	601	32
Sequence 68082, A	US-09-949-016-68082	4	13.7	601	32
Sequence 78126, A	US-09-949-016-78126	4	13.7	601	32
Sequence 79760, A	US-09-949-016-79760	4	13.7	601	32
Sequence 125857, A	US-09-949-016-125857	4	13.7	601	32
Sequence 137324, A	US-09-949-016-137324	4	13.7	601	32
Sequence 145882, A	US-09-949-016-145882	4	13.7	601	32
Sequence 145950, A	US-09-949-016-145950	4	13.7	601	32
Sequence 146218, A	US-09-949-016-146218	4	13.7	601	32
Sequence 152985, A	US-09-949-016-152985	4	13.7	601	32
Sequence 152986, A	US-09-949-016-152986	4	13.7	601	32
Sequence 160719, A	US-09-949-016-160719	4	13.7	601	32
Sequence 160720, A	US-09-949-016-160720	4	13.7	601	32
Sequence 160721, A	US-09-949-016-160721	4	13.7	601	32
Sequence 163504, A	US-09-949-016-163504	4	13.7	601	32
Sequence 163505, A	US-09-949-016-163505	4	13.7	601	32
Sequence 164246, A	US-09-949-016-164246	4	13.7	601	32
Sequence 164318, A	US-09-949-016-164318	4	13.7	601	32
Sequence 172085, A	US-09-949-016-172085	4	13.7	601	32
Sequence 185125, A	US-09-949-016-185125	4	13.7	601	32
Sequence 185126, A	US-09-949-016-185126	4	13.7	601	32
Sequence 185157, A	US-09-949-016-185157	4	13.7	601	32
Sequence 185158, A	US-09-949-016-185158	4	13.7	601	32
Sequence 187447, A	US-09-949-016-187447	4	13.7	601	32
Sequence 205151, A	US-09-949-016-205151	4	13.7	601	32
Sequence 14644, A	US-09-270-767-14644	4	13.7	756	32
Sequence 3645, Ap	US-09-270-767-3645	4	13.7	820	32
Sequence 18927, A	US-09-270-767-18927	4	13.7	820	32
Sequence 239, App	US-08-956-171E-239	4	13.7	1361	32
Sequence 239, App	US-08-781-986A-239	4	13.7	1361	32
Sequence 469, App	US-09-976-594-469	4	13.7	1503	32
Sequence 901, App	US-09-620-312D-901	4	13.7	1732	32
Sequence 231, App	US-09-976-594-231	4	13.7	1953	32
Sequence 5797, Ap	US-09-949-016-5797	4	13.7	2251	32
Sequence 185, App	US-09-311-021-185	4	13.7	2520	32
Sequence 4, Appl1	US-08-100-247-4	1	13.7	2740	32
Sequence 4, Appl1	US-08-483-146A-4	1	13.7	2740	32
Sequence 5, Appl1	US-08-232-513A-5	1	13.7	2740	32
Sequence 4, Appl1	US-08-484-594A-4	1	13.7	2740	32
Sequence 4, Appl1	US-09-076-258A-4	1	13.7	2740	32
Sequence 4, Appl1	US-08-756-031-4	1	13.7	2740	32
Sequence 2, Appl1	US-08-897-443-2	2	13.7	3373	32

c 86 32 13.7 3438 4 US-09-949-016-5648 Sequence 5648, Ap
c 87 32 13.7 3438 4 US-09-949-016-5649 Sequence 5649, Ap
c 88 32 13.7 3449 4 US-09-907-794A-33 Sequence 33, Appl
c 89 32 13.7 3449 4 US-09-905-125A-33 Sequence 33, Appl
c 90 32 13.7 3449 4 US-09-902-775A-33 Sequence 33, Appl
c 91 32 13.7 3449 4 US-09-906-700-33 Sequence 33, Appl
c 92 32 13.7 3449 4 US-09-903-603A-33 Sequence 33, Appl
c 93 32 13.7 3449 4 US-09-904-920A-33 Sequence 33, Appl
c 94 32 13.7 3449 4 US-09-909-064-33 Sequence 33, Appl
c 95 32 13.7 3449 4 US-09-905-381A-33 Sequence 33, Appl
c 96 32 13.7 3449 4 US-09-906-618-33 Sequence 33, Appl
c 97 32 13.7 3496 4 US-09-949-016-344 Sequence 344, App
c 98 32 13.7 4335 4 US-09-949-016-15632 Sequence 15632, A
c 99 32 13.7 10840 4 US-09-949-016-13515 Sequence 13515, A
100 32 13.7 15148 4 US-09-949-016-15058 Sequence 15058, A

ALIGNMENTS

RESULT 1
US-09-949-016-68055/c
; Sequence 68055, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68055
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-68055

Alignment Scores:
Pred. No.: 207 Length: 601
Score: 33.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 14.16% Indels: 0
DB: 4 Gaps: 0

SEQ139-143-144-146-147-17 (1-103) x US-09-949-016-68055 (1-601)

Qy 82 PheTrp***ThrGlu 87
|||||||
Db 407 TTTTGGGACACAG 390

RESULT 2
US-09-949-016-72934/c
; Sequence 72934, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72934
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-72934

Alignment Scores:
Pred. No.: 207 Length: 601
Score: 33.00 Matches: 6
Percent Similarity: 28.00% Conservative: 1
Best Local Similarity: 24.00% Mismatches: 18
Query Match: 14.16% Indels: 0
DB: 4 Gaps: 0

SEQ139-143-144-146-147-17 (1-103) x US-09-949-016-72934 (1-601)

Qy 63 Asp***Leu*****Trp*****Phe 82
|||||
Db 131 GATTCAATCATCTCTTATGGAAGAAGCCCTTGCAATTACCATTTCTTCCACTGATTTT 72

Qy 83 PheTrp***ThrGlu 87
|||||||
Db 71 TTTTGGGGGAGGAG 57

RESULT 3
US-09-949-016-72935/c
; Sequence 72935, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72935
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-72935

Alignment Scores:
Pred. No.: 207 Length: 601
Score: 33.00 Matches: 6
Percent Similarity: 28.00% Conservative: 1
Best Local Similarity: 24.00% Mismatches: 18
Query Match: 14.16% Indels: 0
DB: 4 Gaps: 0

SEQ139-143-144-146-147-17 (1-103) x US-09-949-016-72935 (1-601)

Qy 63 Asp***Leu*****Trp*****Phe 82
|||||
Db 533 GATTCAATCATCTCTTATGGAAGAAGCCCTTGCAATTACCATTTCTTCCACTGATTTT 474

Qy 83 PheTrp***ThrGlu 87
|||||||
Db 473 TTTTGGGGGAGGAG 459

RESULT 4
US-09-949-016-206055/c

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 25, 2005, 22:14:14 ; Search time 577.805 Seconds
(without alignments)
1055.258 Million cell updates/sec

Title: SEQ139-143-144-146-147-17

Perfect score: 233
Sequence: 1 XXXRXXPKXXXRXRXIXXXXX.....WXTXXXXXXXRXRXRXRX 103

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.1
Ygapop 10.0 , Ygapext 0.1
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO spo01/MYERS08974584/runat_25072005_102705_5052/app.query.fasta_1.526
-DB=N_Geneseq_16Dec04 -QFMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomum62 -TRANS=human40.cdi
-LIST=100 -LOCAL=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=MYERS08974584 @CNC 1.1 468 @runat_25072005_102705_5052 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THRAADS=1 -XGAPOP=10 -XGAPEXT=0.1 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.1 -DELOP=6 -DELEXT=7

Database : N_Geneseq_16Dec04:.*
1: Geneseqn1980s:.*
2: Geneseqn1990s:.*
3: Geneseqn2000s:.*
4: Geneseqn2001as:.*
5: Geneseqn2001bs:.*
6: Geneseqn2002as:.*
7: Geneseqn2002bs:.*
8: Geneseqn2003as:.*
9: Geneseqn2003bs:.*
10: Geneseqn2003cs:.*
11: Geneseqn2003ds:.*
12: Geneseqn2004as:.*
13: Geneseqn2004bs:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	35.9	15.4	368	6	ABQ85527 Arabidops
2	35.9	15.4	434	10	ABX61310 Arabidops
3	35.9	15.4	639	6	ABZ14486 Arabidops
4	35	15.0	1794	12	ADJ35133 DNA encod
5	35	15.0	1794	12	ADJ35131 DNA encod

6	35	15.0	2755	10	ADE57777	Ade57777 Human gen
7	35	15.0	2755	10	ADE57781	Ade57781 Human gen
8	34.1	14.6	784	10	ADD30623	Add30623 Plant yie
9	34.1	14.6	784	12	ADI43774	Adi43774 Plant tra
c 10	34.1	14.6	1111	4	AAH29588	Aah29588 Drosophil
11	34	14.6	3761	8	ABZ36127	Abz36127 Human sec
c 12	33	14.2	102	10	ABX06554	Abx06554 S. pneumo
c 13	33	14.2	200	12	ADH00813	Adh00813 Kidney di
14	33	14.2	326	13	ADR63886	Adr63886 Cotton cd
15	33	14.2	376	6	ABN94533	Abn94533 Gene #103
16	33	14.2	422	6	ABN23546	Abn23546 Human ORF
17	33	14.2	488	6	ABV96225	Abv96225 Human pan
18	33	14.2	495	5	ABA19985	Abal9985 Human ner
c 19	33	14.2	572	13	ACN57361	Acn57361 Cotton gy
c 20	33	14.2	586	13	ADQ53130	Adq53130 Novel can
c 21	33	14.2	595	13	ADR63885	Adr63885 Cotton cd
22	33	14.2	649	10	ADD29745	Add29745 Mouse tum
c 23	33	14.2	652	5	ABV52638	Abv52638 Human pro
24	33	14.2	763	4	AAI95301	Aai95301 Human neu
25	33	14.2	876	2	AAI30865	Aax30865 Streptoco
26	33	14.2	889	6	ABQ42238	Abq42238 Oligonucl
c 27	33	14.2	889	6	ABQ42239	Abq42239 Oligonucl
c 28	33	14.2	944	5	AAI57089	Aai57089 DNA encod
29	33	14.2	1110	4	AAH90776	Aah90776 CFE 80 co
30	33	14.2	1113	8	ABZ42222	Abz42222 Streptoco
31	33	14.2	1134	4	AAH90885	Aah90885 2CFE 80 c
32	33	14.2	1171	13	ADS60924	AdS60924 Bacterial
c 33	33	14.2	1174	6	ABQ53835	Abq53835 Oligonucl
34	33	14.2	1174	6	ABQ53834	Abq53834 Oligonucl
35	33	14.2	1311	13	ADQ81928	Adq81928 Human BMP
36	33	14.2	1554	8	ABT15151	Abt15151 Pathogen
37	33	14.2	1575	6	ABN93157	Abn93157 Staphyloc
38	33	14.2	1575	13	ADS04051	AdS04051 Staphyloc
39	33	14.2	1603	6	ABQ60836	Abq60836 FLJ10512f
c 40	33	14.2	1624	2	AAV90773	Aav90773 Nucleotid
41	33	14.2	1650	4	ABL14045	Abi14045 Drosophil
c 42	33	14.2	2226	2	AAV73925	Aav73925 Human SHA
c 43	33	14.2	2269	4	ABL09273	Abi09273 Drosophil
c 44	33	14.2	2269	4	ABL15036	Abi15036 Drosophil
45	33	14.2	2385	4	ABZ13673	Abz13673 Arabidops
46	33	14.2	2418	6	AAH14230	Aah14230 Human cDN
47	33	14.2	2422	4	AAH14230	Aah14230 Human cDN
c 48	33	14.2	2563	2	AAV28617	Aav28617 Nucleotid
c 49	33	14.2	2563	10	ADB68454	AdB68454 Human DCA
c 50	33	14.2	2592	6	ABQ60837	Abq60837 FLJ10512f
c 51	33	14.2	2655	10	ADB68447	AdB68447 Human DCA
c 52	33	14.2	2677	13	ACN37426	Acn37426 Tumour-as
c 53	33	14.2	2677	13	ACN37426	Acn37426 Tumour-as
54	33	14.2	2740	3	AAC36697	Aac36697 Arabidops
55	33	14.2	2741	3	AAC47379	Aac47379 Arabidops
56	33	14.2	2761	4	ABL09154	Abi09154 Drosophil
57	33	14.2	2765	4	ABL28140	Abi28140 Drosophil
c 58	33	14.2	3616	2	AAV73924	Aav73924 Human SHA
c 59	33	14.2	3634	5	ABV23195	Abv23195 Human pro
c 60	33	14.2	3634	5	ABV29032	Abv29032 Human pro
c 61	33	14.2	3986	4	AAF77821	Aaf77821 Murine TS
62	33	14.2	4126	2	AAV90882	Aav90882 Nucleotid
c 63	33	14.2	4292	2	AAV90772	Aav90772 Nucleotid
64	33	14.2	4333	5	ABV28571	Abv28571 Human pro
65	33	14.2	4333	5	ABV22743	Abv22743 Human pro
66	33	14.2	4333	5	ABV27685	Abv27685 Human pro
67	33	14.2	4333	5	ABV21858	Abv21858 Human pro
68	33	14.2	4803	4	ABL09272	Abi09272 Drosophil
69	33	14.2	5201	12	ADP13477	Adp13477 Renal cel
70	33	14.2	5384	6	ABJ32208	Abj32208 Human imm
c 71	33	14.2	6031	5	ABA15148	Abal5148 Human ner
72	33	14.2	6744	4	AAK77797	Aak77797 Human imm
73	33	14.2	6915	6	ABL32406	Abi32406 Human imm
74	33	14.2	7071	13	ADR84283	Adr84283 Aspergill
75	33	14.2	18627	2	AAV52246	Aav52246 Streptoco
c 76	33	14.2	4096	10	ADB68445	AdB68445 Human DCA
c 77	33	14.2	88853	11	ACN43904	Acn43904 Mouse gen
78	33	14.2	110000	10	ABS56454_07	Continuation (8 of

79 33 14.2 110000 12 ADN46845_10 Continuation (11 o
 C 80 33 14.2 110000 12 ADN47591_09 Continuation (11 o
 C 81 33 14.2 110000 12 ADN47591_10 Continuation (11 o
 C 82 33 14.2 110000 12 ADN46123_10 Continuation (11 o
 C 83 33 14.2 110000 12 ADN47209_09 Continuation (11 o
 C 84 33 14.2 110000 12 ADN47209_10 Continuation (11 o
 C 85 33 14.2 110000 12 ADN46464_10 Continuation (11 o
 C 86 33 14.2 110000 12 ADN47960_09 Continuation (11 o
 C 87 33 14.2 110000 12 ADN47960_10 Continuation (11 o
 C 88 33 14.2 110000 13 ABD32966_02 Continuation (3 of
 C 89 33 14.2 110000 13 ABD32966_03 Continuation (4 of
 C 90 33 14.2 110000 13 ABD32806_14 Continuation (5 of
 C 91 33 14.2 179651 10 ADL13813 Adl13813 Osteoarth
 C 92 33 14.2 209484 11 ACN44126 Acn44126 Human gen
 C 93 33 14.2 210204 12 ADQ18927 Adq18927 Human sof
 C 94 33 14.2 231222 10 ADL13693 Adl13693 Osteoarth
 C 95 32.9 14.1 568 13 ACN45985 Acn45985 Cotton pr
 C 96 32.8 14.1 237 8 ABZ78105 Abz78105 Human Sup
 C 97 32.8 14.1 419 10 ADI02612 Adi02612 Human cDN
 C 98 32.8 14.1 458 3 AAC16991 Aac16991 Human sec
 C 99 32.8 14.1 476 10 ADB99053 Adb99053 Lrp5 cons
 C 100 32.8 14.1 476 10 ADE82622 Ade82622 pcdNA3.1

ALIGNMENTS

RESULT 1
 ABQ85527
 ID ABQ85527 standard; DNA; 368 BP.
 XX
 AC ABQ85527;
 XX
 DT 05-SEP-2002 (first entry)
 XX
 DE Arabidopsis thaliana expressed polynucleotide SEQ ID NO 397.
 XX
 KW Plant; Arabidopsis; transgenic; fungicide; insecticide; ds.
 XX
 OS Arabidopsis thaliana.
 XX
 PN US2002062014-A1.
 XX
 PD 23-MAY-2002.
 XX
 PF 26-JAN-2001; 2001US-00770791.
 XX
 PR 27-JAN-2000; 2000US-0178480P.
 XX
 PA (GORL/) GORLACH J.
 PA (ANY/) AN Y.
 PA (HAMI/) HAMILTON C M.
 PA (PRIC/) PRICE J L.
 PA (RAIN/) RAINES T M.
 PA (YUY/) YU Y.
 PA (RAME/) RAMEAKA J G.
 PA (PAGE/) PAGE A.
 PA (MATH/) MATHAW A V.
 PA (LEDF/) LEDFORD B L.
 PA (WOES/) WOESSNER J P.
 PA (HAAS/) HAAS W D.
 PA (GARC/) GARCIA C A.
 PA (KRIC/) KRICKER M.
 PA (SLAT/) SLATER T.
 PA (DAVI/) DAVIS K R.
 PA (ALLE/) ALLEN K.
 PA (HOFF/) HOFFMAN N.
 PA (HURB/) HURBAN P.
 XX
 PI Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;
 PI Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
 PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;
 PI Hurban P;
 XX

DR WPI; 2002-479265/51.
 XX New nucleic acid sequences of Arabidopsis thaliana and their encoded products are useful to produce transgenic plants, to screen for biologically active agents such as fungicides and insecticides and in genetic studies.
 PT
 PT Claim 1; SEQ ID NO 397; 18pp + Sequence Listing; English.
 XX
 PS The invention relates to a novel nucleic acid of Arabidopsis thaliana comprising a sequence capable of hybridising under stringency to one of the 99 sequences referred to but not defined in the specification (ABQ85131-ABQ86129). The nucleic acid sequences are useful to identify homologous or related genes, to produce compositions that modulate expression or function of the encoded protein, to map functional regions of the protein, to study associated physiological pathways, to genetically manipulate cells and plants. The encoded products are useful to screen for biologically active agents such as fungicides or insecticides and to elucidate biochemical pathways
 CC
 CC Sequence 368 BP; 87 A; 100 C; 105 G; 73 T; 0 U; 3 Other;
 SQ

Alignment Scores:
 Pred. No.: 106 Length: 368
 Score: 35.90 Matches: 7
 Percent Similarity: 26.92% Conservative: 0
 Best Local Similarity: 26.92% Mismatches: 8
 Query Match: 15.41% Indels: 11
 DB: 6 Gaps: 1

SEQ139-143-144-146-147-17 (1-103) x ABQ85527 (1-368)

Qy 62 AspAsp***Leu*****Trrp***** 81
 |||||
 Db 231 GACGACTCTATCTGCACGTGG-----CGG 257

Qy 82 PhePheTrrp***ThrGlu 87
 |||||
 Db 258 CTCCTGTGGACGACGGAG 275

RESULT 2
 ABX61310
 ID ABX61310 standard; DNA; 434 BP.
 XX
 AC ABX61310;
 XX
 DT 26-FEB-2003 (first entry)
 XX
 DE Arabidopsis thaliana polynucleotide #656.
 XX
 KW Thale cress; gene; ds; genetic manipulation; plant; biosynthesis;
 KW genetic modification; environmental stress; disease resistance;
 KW fungicide; insecticide; stress tolerance.
 XX
 OS Arabidopsis thaliana.
 XX
 PN US2002142319-A1.
 XX
 PD 03-OCT-2002.
 XX
 PF 07-AUG-2001; 2001US-00924035.
 XX
 PR 13-AUG-1999; 99US-0148784P.
 PR 11-AUG-2000; 2000US-00638258.
 XX
 PA (GORL/) GORLACH J.
 PA (ANY/) AN Y.
 PA (HAMI/) HAMILTON C M.
 PA (PRIC/) PRICE J L.
 PA (HARG/) HARGISS T R.
 PA (YUY/) YU Y.
 PA (RAME/) RAMEAKA J G.
 PA (PAGE/) PAGE A.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 25, 2005, 22:14:49 ; Search time 3386.44 Seconds
(without alignments)
1473.786 Million cell updates/sec

Title: SEQ139-143-144-146-147-17

Perfect score: 233

Sequence: 1 XXXRXXPKXXXRXIXXXXXX.....WXTXXXXXXXRXRXRXRX 103

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.1
Ygapop 10.0, Ygapext 0.1
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cpn2.1/USPTO.spool/MYERS08974584/runat_25072005.102706.5062/app.query.fasta_1.526
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEX=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=MYERS08974584 @CN 1.1.2647 @runat_25072005.102706.5062 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.1 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.1 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sv.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
C 1	41.9	18.0	336	8	AY201359 Arabidops
C 2	39.9	17.1	185084	5	BX629350 Zebrafish
C 3	39.9	17.1	202699	2	BX640481 Danio rer
C 4	39	16.7	104128	10	AL929562 Mouse DNA

5	38	16.3	245146	2	AC130643	AC130643	Rattus no	
c	6	38	16.3	253634	2	AC115167	Rattus no	
7	37	15.9	106208	2	AC138061	AC138061	Homo sapi	
8	37	15.9	175245	10	AC129188	AC129188	Mus muscu	
9	37	15.9	175542	2	AC150062	AC150062	Gallus ga	
10	37	15.9	187731	10	AC140307	AC140307	Mus muscu	
c	11	37	15.9	188771	2	AC150073	Gallus ga	
c	12	36.8	15.8	37841	2	AC017708	Drosophill	
c	13	36.8	15.8	168471	3	AC006170	Drosophill	
c	14	36.8	15.8	231562	3	AE003767	Drosophill	
15	36.1	15.5	4320	10	BC060679	BC060679	Mus muscu	
16	36	15.5	110000	2	AC096393_5	Continuation (6 of		
17	36	15.5	181561	2	AC104098	AC104098	Mus muscu	
18	36	15.5	215745	10	AC117585	AC117585	Mus muscu	
19	36	15.5	231612	2	AC130088	AC130088	Rattus no	
20	36	15.5	245044	2	AC131471	AC131471	Rattus no	
21	36	15.5	271485	2	AC137162	AC137162	Rattus no	
c	22	36	15.5	281085	2	AC123455	Rattus no	
23	35.9	15.4	573	8	AY201051	AY201051	Arabidops	
24	35.9	15.4	591	8	BT015452	BT015452	Arabidops	
25	35.9	15.4	639	6	AX507596	AX507596	Sequence	
26	35.9	15.4	639	8	AY074558	AY074558	Arabidops	
27	35.9	15.4	702	8	AY054226	AY054226	Arabidops	
28	35.9	15.4	2259	8	BT001993	BT001993	Arabidops	
29	35.9	15.4	95643	8	FN6N15	AF069299	Arabidops	
c	30	35.9	15.4	174269	2	AC132961	Rattus no	
c	31	35.9	15.4	197119	8	ATCHRIV1	Arabidops	
c	32	35.9	15.4	199709	2	AC113049	Mus muscu	
33	35.9	15.4	206887	2	AC134007	AC134007	Rattus no	
34	35.9	15.4	259970	2	AC106985	AC106985	Rattus no	
35	35.1	15.1	639	8	AF528583	AF528583	Arabidops	
36	35.1	15.1	222348	10	AC111103	AC111103	Mus muscu	
37	35	15.0	1611	8	AK067539	AK067539	Oryza sat	
c	38	35	15.0	127614	8	AP005879	Oryza sat	
c	39	35	15.0	154912	8	AP005684	Oryza sat	
c	40	35	15.0	231400	2	AC109895	Rattus no	
41	34.9	15.0	705	8	AY203308	AY203308	Arabidops	
c	42	34.9	15.0	218729	2	AC111863	Rattus no	
c	43	34.6	14.8	110000	2	AC151898_4	Continuation (5 of	
c	44	34.6	14.8	110000	2	AC151898_5	Continuation (6 of	
c	45	34.6	14.8	153108	10	AC142450	Mus muscu	
c	46	34.3	14.7	237624	2	AC113662	Rattus no	
47	34.3	14.7	247917	2	AC095906	AC095906	Rattus no	
48	34.1	14.6	639	8	AF528584	AF528584	Arabidops	
49	34.1	14.6	639	8	AF528586	AF528586	Arabidops	
50	34.1	14.6	639	8	AF528588	AF528588	Arabidops	
51	34.1	14.6	639	8	AF528589	AF528589	Arabidops	
52	34.1	14.6	639	8	AF528590	AF528590	Arabidops	
53	34.1	14.6	681	8	AY141230	AY141230	Arabidops	
c	54	34.1	14.6	1111	6	AX094653	Sequence	
c	55	34.1	14.6	25749	2	AC014588	AC014588	Drosophill
56	34.1	14.6	88292	8	T8F5	AC004512	Arabidops	
c	57	34.1	14.6	175519	3	AC010666	AC010666	Drosophill
c	58	34.1	14.6	250029	3	AE003599	AE003599	Drosophill
c	59	34	14.6	46189	8	AP006676	AP006676	Lotus cor
60	34	14.6	147109	9	AL358779	AL358779	Human DNA	
61	34	14.6	149464	10	AC137096	AC137096	Mus muscu	
62	34	14.6	159119	1	AP006583	AP006583	Gloeobact	
63	34	14.6	159952	9	AC007252	AC007252	Homo sapi	
64	34	14.6	160421	2	AC025005	AC025005	Homo sapi	
c	65	34	14.6	173062	9	AL445587	AL445587	Human DNA
c	66	34	14.6	175639	10	AC133515	AC133515	Mus muscu
c	67	34	14.6	177675	10	AC124534	AC124534	Mus muscu
c	68	34	14.6	181731	2	AC102525	Mus muscu	
c	69	34	14.6	183772	9	AC021006	AC021006	Homo sapi
c	70	34	14.6	206444	2	AC113538	AC113538	Mus muscu
c	71	34	14.6	217330	10	AC118414	AC118414	Rattus no
c	72	34	14.6	221697	2	AC102395	AC102395	Mus muscu
73	34	14.6	223437	2	AC110660	AC110660	Rattus no	
74	34	14.6	237289	2	AC095523	AC095523	Rattus no	
75	34	14.6	238352	2	AC114102	AC114102	Rattus no	
76	34	14.6	238749	2	AC110965	AC110965	Rattus no	
77	34	14.6	240123	2	AC098494	AC098494	Rattus no	

```

c 78 33.9 14.5 28054 3 CEY4081A
79 33.8 14.5 216813 2 AC148423
80 33.8 14.5 240857 2 AC148371
81 33.1 14.2 152748 9 AL1354777
82 33.1 14.2 155539 9 AL138721
c 83 33 14.2 102 6 AX568476
c 84 33 14.2 248 9 HG050555
c 85 33 14.2 282 6 CQ742132
c 86 33 14.2 316 3 AY427333
c 87 33 14.2 320 3 AY427365
c 88 33 14.2 336 6 CQ745819
c 89 33 14.2 376 6 AX408384
90 33 14.2 422 6 CQ449809
91 33 14.2 494 11 G31161
92 33 14.2 620 9 AF29307782
c 93 33 14.2 652 6 CQ520790
94 33 14.2 665 10 BC028521
95 33 14.2 763 6 BD019138
96 33 14.2 763 6 BD099076
c 97 33 14.2 947 9 BC042148
c 98 33 14.2 1080 9 HSDGCR6
99 33 14.2 1110 6 AX194095
100 33 14.2 1134 6 AX194317

```

ALIGNMENTS

```

RESULT 1
AY201359/c 336 bp DNA linear PLN 24-JAN-2003
LOCUS Arabidopsis thaliana sequence flanking Ds5 end of Ds-Trap insertion
DEFINITION from line GT260.
ACCESSION AY201359
VERSION AY201359.1 GI:27897313
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicaceae; Arabidopsis.
1 (bases 1 to 336)
REFERENCE May,B.P., Simorowski,J., Arroyo,J.-M., Vaughn,M.W., Shen,R.,
AUTHORS McCombie,W.R. and Martienssen,R.A.
Arabidopsis genomic sequences flanking Ds enhancer and gene traps
in transgenic lines
TITLE Arabidopsis thaliana (thale cress)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 336)
AUTHORS May,B.P., Simorowski,J., Arroyo,J.-M., Vaughn,M.W., Shen,R.,
McCombie,W.R. and Martienssen,R.A.
Direct Submission
TITLE Arabidopsis thaliana
JOURNAL Submitted (17-DEC-2002) Cold Spring Harbor Laboratory, 1 Bungtown
Road, Cold Spring Harbor, NY 11724, USA
COMMENT http://genetrapp.cshl.org.
FEATURES Location/Qualifiers
source 1..336
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/clone="GT260.Ds5.04.01.99.b.336"
/ecotype="Landsberg"
/note="transgenic line GT260"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 0.453 Length: 336
Score: 41.90 Matches: 8
Percent Similarity: 30.77% Conservative: 0
Best Local Similarity: 30.77% Mismatches: 7
Query Match: 17.98% Indels: 11
DB: 8 Gaps: 1

```

```

SEQ139-143-144-146-147-17 (1-103) x AY201359 (1-336)

```

```

Oy 62 AspAsp***Leu*****Trr*****
Db 235 GACGACTCTCTATCTGCTAGCTGG-----CGG 209
Oy 82 PhePheTrp***ThrGlu 87
Db 208 TTCTTGTGGACGACGGAG 191
RESULT 2
BX629350/c 185084 bp DNA linear VRT 19-MAR-2004
LOCUS Zebrafish DNA sequence from clone CH211-155E13 in linkage group 10,
DEFINITION complete sequence.
ACCESSION BX629350
VERSION BX629350.6 GI:45581081
KEYWORDS HTG.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 185084)
REFERENCE Henderson,C.
AUTHORS Direct Submission
TITLE Submitted (19-MAR-2004) Wellcome Trust Sanger Institute, Hinxton,
JOURNAL Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Mar 16, 2004 this sequence version replaced gi:45379234.
COMMENT ----- Genome Center
Center: Wellcome Trust Sanger Institute
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
-----
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/Celegans/wormpep/Clone-derived
Zebrafish pUC subclones occasionally display inconsistency over the
length of mononucleotide A/T runs and conserved TA repeats. Where
this is found the longest good quality representation will be
submitted.
Repeat names beginning 'Dr' were identified by the Recon repeat
discovery system (Zhairong Bao and Sean Eddy, submitted), and those
beginning 'drr' were identified by Rick Waterman (Stephen Johnson
lab, WashU). For further information see
http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml
CH211-155E13 is from a CHORI-211 BAC library
VECTOR: pIARBAC2.1.
Location/Qualifiers
1..185084
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-155E13"
/clone_lib="CHORI-211"

```

FEATURES

```

source

```

ORIGIN

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 25, 2005, 22:18:34 ; Search time 4138.71 Seconds

(without alignments)
938.107 Million cell updates/sec

Title: SEQ139-143-144-146-147-16

Perfect score: 232

Sequence: 1 XXXXXPPXXXXXIXXXXX.....FWTXXXXXXXXXXXXXX 102

Scoring table:

BLASOP62
Xgapop 10.0 , Xgapext 0.1
Ygapop 10.0 , Ygapext 0.1
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO spoil/MYERS08974584/runat_25072005_102706_5075/app_query.fasta_1.526
-DB=EST -QFMT=fastap -SUFFIX=first -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blasop62 -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HRAPIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=MYERS08974584@cgn.1.1.3556 @runat_25072005_102706_5075 -NCPU=6 -ICPU=3
-NO.MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN TIMEOUT=30 -THRAIDS=1 -XGAPOP=10 -XGAPEXT=0.1 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.1 -DELOP=6 -DELEXT=7

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gse1: *
9: gb_gse2: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query	ID	Description
C 1	39	16.8	352	5	BP427888
C 2	39	16.8	437	5	BUS06243
C 3	39	16.8	576	5	BP760208
4	39	16.8	599	4	BI964368
5	39	16.8	601	5	BP766928
6	39	16.8	642	5	BP771659
7	39	16.8	673	9	CL186598
8	39	16.8	707	6	CD741981
9	38	16.4	895	2	BEG17363
					BP427888 BP427888
					BUS06243 AGENCOURT
					BP760208 BP760208
					BI964368 id38a04.x
					BP766928 BP766928
					BP771659 BP771659
					CL186598 104_401.1
					CD741981 UI-M-AOO-
					BEG17363 601441886

10	38	16.4	1047	5	BQ224762
11	37	15.9	939	9	CL398497
C 12	36.8	15.9	851	5	BUS09508
13	36.4	15.7	1058	8	BH725011
14	36.2	15.6	181	9	CG784907
15	36.2	15.6	197	9	CG887155
16	36.2	15.6	216	6	CB178950
17	36.2	15.6	357	9	CG784906
C 18	36.2	15.6	408	7	CK225321
19	36.2	15.6	563	7	CK225321
20	36.2	15.6	578	7	CN717076
21	36.2	15.6	624	7	CN699623
22	36.2	15.6	665	2	BB616800
23	36.2	15.6	741	7	CF733575
24	36.2	15.6	746	7	CK780407
25	36.2	15.6	748	6	CD803163
26	36.2	15.6	768	7	CO433173
27	36.2	15.6	859	7	CO807249
28	36.2	15.6	3232	3	AK030089
29	36.2	15.6	4003	3	AK028324
30	36.2	15.6	4024	3	AK030951
31	36.2	15.6	4306	3	AK032405
32	36.2	15.6	4942	3	AK031688
33	35.9	15.5	548	7	W43412
34	35.9	15.5	598	5	BF561807
35	35.9	15.5	651	3	CNS0A2QM
36	35.9	15.5	680	5	BX836890
37	35.9	15.5	681	1	AV822745
38	35.9	15.5	692	3	CNS0A2P9
39	35.9	15.5	693	3	CNS0YXGP
C 40	35.9	15.5	753	8	AZ869871
41	35.9	15.5	758	3	CNS0A2LE
42	35.9	15.5	893	7	CF652848
43	35.9	15.5	900	7	CV242818
C 44	35.9	15.5	928	9	CL514461
C 45	35.3	15.2	563	9	AX9889710
46	35.1	15.1	582	1	AJ799564
47	35.1	15.1	716	1	AJ560032
C 48	35.1	15.1	1107	5	BQ618909
C 49	35.1	15.1	1107	5	BQ619092
C 50	35.1	15.1	1107	5	BQ619227
C 51	35.1	15.1	1107	5	BQ619307
C 52	35.1	15.1	1107	5	BQ619325
53	35	15.1	591	4	BG840829
54	35	15.1	669	8	AQ840198
55	35	15.1	761	4	BG839900
56	35	15.1	794	7	CK305937
57	34.9	15.0	449	8	AZ216917
C 58	34.9	15.0	554	9	TA271D10Q
C 59	34.2	14.7	648	8	CB461221
C 60	34.1	14.7	206	8	BZ660880
61	34	14.7	148	1	AA514345
62	34	14.7	229	1	AI335300
C 63	34	14.7	239	1	AI335291
64	34	14.7	263	7	H09046
65	34	14.7	278	5	BW171513
66	34	14.7	278	5	BW171535
C 67	34	14.7	294	5	BW300646
C 68	34	14.7	297	2	AW303078
69	34	14.7	323	7	H18423
70	34	14.7	365	2	BF395855
71	34	14.7	381	1	AA969353
C 72	34	14.7	409	5	BM923188
73	34	14.7	416	2	BF920885
74	34	14.7	420	1	AI635438
75	34	14.7	424	8	AQ084051
76	34	14.7	429	2	BF409184
C 77	34	14.7	450	5	EX480972
78	34	14.7	450	8	A2129520
79	34	14.7	459	1	A274649
80	34	14.7	462	1	AA861318
81	34	14.7	462	1	AJ003578
82	34	14.7	493	6	CA432939
					BQ224762 AGENCOURT
					CL398497 ZMMBB039
					BUS09508 AGENCOURT
					BH725011 BOMNB89TF
					CG784907 RRR157 Ba
					CG887155 RRS688 Ba
					CB178950 in99B04.y
					CG784906 RRR158 Ba
					CK225321 703207690
					CN722398 E0841A10-
					CN717076 E0737G08-
					CN699623 E0423H10-
					BB616800 BB616800
					CF733575 UI-M-HB0-
					CK780407 UI-M-GV0-
					CD803163 UI-M-GV0-
					CO433173 UI-M-HX0-
					CO807249 AGENCOURT
					AK030089 Mus muscu
					AK028324 Mus muscu
					AK030951 Mus muscu
					AK032405 Mus muscu
					AK031688 Mus muscu
					W43412 22805 CD4-1
					BF561807 BF561807
					BX836890 BX836890
					AV822745 AV822745
					AX988971 2M0182H11
					AX841609 Arabidops
					AX829061 Arabidops
					AX829061 Arabidops
					AZ869871 2M0182H11
					CF652848 80-L02052
					CV242818 WS02516.B
					CL514461 SAIL 889
					AX9889710 Forward 5
					AJ799564 AJ799564
					AJ560032 AJ560032
					BQ618909 RNSQ1C0
					BQ619092 RNSQ1C0
					BQ619227 RNSQ1C0
					BQ619307 RNSQ1C0
					BQ619325 RNSQ1C0
					BG840829 MEST12-D0
					AQ840198 rxb0052M
					BG839900 Gm01_09b1
					CK305937 SB0303182
					AZ216917 Sheared D
					AL488262 T. brucei
					CB461221 tigr-g89
					BZ660880 SALK 0243
					AA514345 nf56D04.B
					AI335300 ta47b09.x
					AI335291 ta47a09.x
					H09046 y196d12.s1
					BW171513 BW171513
					BW171535 BW171535
					BW300646 BW300646
					AW303078 xr8e07.x
					H18423 ym43a01.s1
					BF395855 UI-R-B82-
					AA969353 on58c03.B
					BM923188 AGENCOURT
					BF920885 MR2-NT013
					AI635438 t858g06.x
					AQ084051 HS_2230.B
					BF409184 UI-R-BT1-
					EX480972 DKFZp686A
					A2129520 OSUNB010
					A274649 ai1d01.B
					AA861318 ak14h10.B
					AJ003578 ai88d05.B
					CA432939 UI-H-C00-

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 26, 2005, 00:55:25 ; Search time 1176.73 Seconds
(without alignments)
560.502 Million cell updates/sec

Title: SEQ139-143-144-146-147-16
Perfect score: 232
Sequence: 1 XXXRXXPKXXXXXIXXXXX.....FWXTXXXXXXXRXKXW 102

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.1
Ygapop 10.0, Ygapext 0.1
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 7277826 seqs, 3233139505 residues

Total number of hits satisfying chosen parameters: 14555652

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters: -MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/MyERS08974584/runat_25072005_102708_5180/app.query.fasta_1.526
-DB=Published_Applications_NA -QFM=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=100 -LOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=40 -MODE=BLAST -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000
-USER=MYERS08974584 @CEN 1.1.456 @runat_25072005_102708_5180 -NCPU=6 -ICPU=3
-NO MWAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=10 -XGAPOP=10 -XGAPEXT=0.1 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.1 -DELOP=6 -DELEXT=7

Database : Published Applications NA.*
1: /cgn2_6/ptodata/2/pubna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/2/pubna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35.9	15.5	368	9	US-09-770-791-397 Sequence 397, App
2	35.9	15.5	434	9	US-09-924-035A-656 Sequence 656, App
3	35.9	15.5	639	9	US-09-938-842A-2291 Sequence 2291, Ap
4	35.9	15.5	639	11	US-09-938-842A-2291 Sequence 2291, Ap
5	35.9	15.5	2321	20	US-10-739-930-1097 Sequence 1097, Ap
6	35	15.1	455	19	US-10-437-963-9511 Sequence 9511, Ap
7	35	15.1	600	22	US-10-972-079-62547 Sequence 62547, A
8	35	15.1	829	20	US-10-425-115-42905 Sequence 42905, A
9	34.1	14.7	784	17	US-10-225-066A-655 Sequence 655, App
10	34.1	14.7	784	17	US-10-374-780A-2237 Sequence 2237, Ap
11	34	14.7	404	18	US-10-424-599-40084 Sequence 40084, A
12	34	14.7	540	22	US-10-972-079-22314 Sequence 22314, A
13	34	14.7	1002	18	US-10-424-599-102834 Sequence 102834, A
14	34	14.7	1383	19	US-10-437-963-100022 Sequence 100022, A
15	33	14.2	102	21	US-10-472-928-1683 Sequence 1683, Ap
16	33	14.2	140	9	US-09-783-590-278 Sequence 278, App
17	33	14.2	193	19	US-10-437-963-102089 Sequence 102089, A
18	33	14.2	201	20	US-10-719-993-25917 Sequence 25917, A
19	33	14.2	282	18	US-10-424-599-57538 Sequence 57538, A
20	33	14.2	326	19	US-10-767-795-4667 Sequence 4667, Ap
21	33	14.2	330	9	US-09-783-590-11246 Sequence 11246, A
22	33	14.2	376	9	US-09-880-107-1031 Sequence 1031, Ap
23	33	14.2	460	20	US-10-425-115-80762 Sequence 80762, A
24	33	14.2	488	14	US-10-060-036-1633 Sequence 1633, Ap
25	33	14.2	521	20	US-10-425-115-36993 Sequence 36993, A
26	33	14.2	529	19	US-10-767-701-5515 Sequence 5515, Ap
27	33	14.2	539	18	US-10-424-599-29238 Sequence 29238, A
28	33	14.2	572	19	US-10-021-323-12142 Sequence 12142, A
29	33	14.2	593	17	US-10-027-632-271755 Sequence 271755, A
30	33	14.2	593	17	US-10-027-632-271755 Sequence 271755, A
31	33	14.2	595	13	US-10-767-795-4666 Sequence 4666, Ap
32	33	14.2	616	13	US-10-027-632-265805 Sequence 265805, A
33	33	14.2	616	13	US-10-027-632-265806 Sequence 265806, A
34	33	14.2	616	17	US-10-027-632-265805 Sequence 265805, A
35	33	14.2	616	17	US-10-027-632-265806 Sequence 265806, A
36	33	14.2	619	13	US-10-027-632-196891 Sequence 196891, A
37	33	14.2	619	13	US-10-027-632-196892 Sequence 196892, A
38	33	14.2	619	13	US-10-027-632-196893 Sequence 196893, A
39	33	14.2	619	17	US-10-027-632-196891 Sequence 196891, A
40	33	14.2	619	17	US-10-027-632-196892 Sequence 196892, A
41	33	14.2	619	17	US-10-027-632-196893 Sequence 196893, A
42	33	14.2	649	20	US-10-335-053-194 Sequence 194, App
43	33	14.2	652	20	US-10-357-930-52657 Sequence 52657, A
44	33	14.2	664	18	US-10-025-114-16192 Sequence 16192, A
45	33	14.2	667	13	US-10-027-632-287415 Sequence 287415, A
46	33	14.2	667	13	US-10-027-632-287416 Sequence 287416, A
47	33	14.2	667	17	US-10-027-632-287415 Sequence 287415, A
48	33	14.2	667	17	US-10-027-632-287416 Sequence 287416, A
49	33	14.2	721	13	US-10-027-632-260266 Sequence 260266, A
50	33	14.2	721	17	US-10-027-632-260266 Sequence 260266, A
51	33	14.2	900	20	US-10-425-115-160420 Sequence 160420, A
52	33	14.2	987	18	US-10-424-599-90405 Sequence 90405, A
53	33	14.2	1113	19	US-10-474-776-76 Sequence 76, Appl
54	33	14.2	1162	21	US-10-956-157-1455 Sequence 1455, Ap
55	33	14.2	1162	17	US-10-956-157-6690 Sequence 6690, Ap
56	33	14.2	1171	17	US-10-369-493-36598 Sequence 36598, A
57	33	14.2	1171	20	US-10-363-345A-40425 Sequence 40425, A
58	33	14.2	1174	20	US-10-363-345A-40426 Sequence 40426, A
59	33	14.2	1174	21	US-10-363-483A-40425 Sequence 40425, A
60	33	14.2	1174	21	US-10-363-483A-40426 Sequence 40426, A
61	33	14.2	1201	20	US-10-425-115-89303 Sequence 89303, A
62	33	14.2	1312	18	US-10-425-114-25327 Sequence 25327, A
63	33	14.2	1404	20	US-10-425-115-49114 Sequence 49114, A
64	33	14.2	1554	21	US-10-470-048B-537 Sequence 537, App

65 33 14.2 1575 22 US-10-724-972A-3346
c 66 33 14.2 1631 13 US-10-027-632-258931
c 67 33 14.2 1631 13 US-10-027-632-258932
c 68 33 14.2 1631 13 US-10-027-632-258933
c 69 33 14.2 1631 13 US-10-027-632-258934
c 70 33 14.2 1631 17 US-10-027-632-258931
c 71 33 14.2 1631 17 US-10-027-632-258932
c 72 33 14.2 1631 17 US-10-027-632-258933
c 73 33 14.2 1631 17 US-10-027-632-258934
c 74 33 14.2 1988 13 US-10-027-632-97163
c 75 33 14.2 1988 13 US-10-027-632-97164
c 76 33 14.2 1988 13 US-10-027-632-97165
c 77 33 14.2 1988 17 US-10-027-632-97163
c 78 33 14.2 1988 17 US-10-027-632-97164
c 79 33 14.2 1988 17 US-10-027-632-97165
c 80 33 14.2 2418 9 US-09-938-842A-1478
c 81 33 14.2 2418 11 US-09-938-842A-1478
c 82 33 14.2 2563 9 US-09-782-051-1
c 83 33 14.2 2655 21 US-10-876-285-3
c 84 33 14.2 2760 20 US-10-739-930-157
c 85 33 14.2 3634 20 US-10-357-930-23184
c 86 33 14.2 3634 20 US-10-357-930-29050
c 87 33 14.2 3986 13 US-10-092-925-1
c 88 33 14.2 4105 17 US-10-062-674-1991
c 89 33 14.2 4333 20 US-10-357-930-21849
c 90 33 14.2 4333 20 US-10-357-930-22736
c 91 33 14.2 4333 20 US-10-357-930-27699
c 92 33 14.2 4333 20 US-10-357-930-28584
c 93 33 14.2 5201 19 US-10-717-597-213
c 94 33 14.2 5384 15 US-10-311-455-181
c 95 33 14.2 6915 15 US-10-311-455-379
c 96 33 14.2 18627 8 US-08-961-527-113
c 97 33 14.2 18627 17 US-10-158-844-113
c 98 33 14.2 44096 21 US-10-876-285-1
c 99 33 14.2 88853 13 US-10-087-192-85
c 100 33 14.2 127602 22 US-10-737-082-65

ALIGNMENTS

RESULT 1
US-09-770-791-397
; Sequence 397, Application US/0970791
; Patent No. US20020062014A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Kricker, Carlos A.
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2029 (PARA-018PRV)
; CURRENT APPLICATION NUMBER: US/09/770,791
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,480
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 397
; LENGTH: 368
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(368)
; OTHER INFORMATION: n = A,T,C or G
US-09-770-791-397
Alignment Scores:
Pred. No.: 183 Length: 368
Score: 35.90 Matches: 7
Percent Similarity: 26.92% Conservative: 0
Best Local Similarity: 26.92% Mismatches: 8
Query Match: 15.47% Indels: 11
DB: 9 Gaps: 1
SEQ139-143-144-146-147-16 (1-102) x US-09-770-791-397 (1-368)
Qy 62 Asplasp***Leu*****Trp***** 81
Db 231 GACGACTCTCTATCTGCCACGTGG-----CGG 257
Qy 82 PhePheTrp***ThrGlu 87
Db 258 CTCTTGTGGACGACGGAG 275
RESULT 2
US-09-924-035A-656
; Sequence 656, Application US/09924035A
; Patent No. US20020142319A1
; GENERAL INFORMATION:
; APPLICANT: Grisch, Jrn
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2011US
; CURRENT APPLICATION NUMBER: US/09/924,035A
; CURRENT FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: US 60/148,784
; PRIOR FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 900
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 656
; LENGTH: 434
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(434)
; OTHER INFORMATION: n = A,T,C or G
US-09-924-035A-656
Alignment Scores:
Pred. No.: 216 Length: 434
Score: 35.90 Matches: 7
Percent Similarity: 26.92% Conservative: 0
Best Local Similarity: 26.92% Mismatches: 8
Query Match: 15.47% Indels: 11
DB: 9 Gaps: 1
SEQ139-143-144-146-147-16 (1-102) x US-09-924-035A-656 (1-434)
Qy 62 Asplasp***Leu*****Trp***** 81
Db 293 GACGACTCTCTATCTGCCACGTGG-----CGG 319
Qy 82 PhePheTrp***ThrGlu 87
Db 320 CTCTTGTGGACGACGGAG 337
RESULT 3
US-09-938-842A-2291

Sequence 2291, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: SCRI1300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
PRIOR FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 2291
LENGTH: 639
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-2291

Alignment Scores:
Pred. No.: 322 Length: 639
Score: 35.90 Matches: 7
Percent Similarity: 26.92% Conservative: 0
Best Local Similarity: 26.92% Mismatches: 8
Query Match: 15.47% Indels: 11
DB: 9 Gaps: 1

SEQ139-143-144-146-147-16 (1-102) x US-09-938-842A-2291 (1-639)

Qy 62 AspAsp***Leu*****Trp***** 81
Db 241 GACGACTCTCTATCTGCCACGTGG-----CGG 267

Qy 82 PhePheTrp***ThrGlu 87
Db 268 CTCCTGTGGACGACGGAG 285

RESULT 4
US-09-938-842A-2291
Sequence 2291, Application US/09938842A
Publication No. US20040009476A9
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: SCRI1300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 2291
LENGTH: 639
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-2291

Alignment Scores:
Pred. No.: 322 Length: 639
Score: 35.90 Matches: 7
Percent Similarity: 26.92% Conservative: 0
Best Local Similarity: 26.92% Mismatches: 8
Query Match: 15.47% Indels: 11
DB: 9 Gaps: 1

Percent Similarity: 26.92% Conservative: 0
Best Local Similarity: 26.92% Mismatches: 8
Query Match: 15.47% Indels: 11
DB: 11 Gaps: 1

SEQ139-143-144-146-147-16 (1-102) x US-09-938-842A-2291 (1-639)

Qy 62 AspAsp***Leu*****Trp***** 81
Db 241 GACGACTCTCTATCTGCCACGTGG-----CGG 267

Qy 82 PhePheTrp***ThrGlu 87
Db 268 CTCCTGTGGACGACGGAG 285

RESULT 5
US-10-739-930-1097
Sequence 1097, Application US/10739930
Publication No. US20040216190A1
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
FILE REFERENCE: 38-21(53377)B
CURRENT APPLICATION NUMBER: US/10/739,930
CURRENT FILING DATE: 2003-12-18
NUMBER OF SEQ ID NOS: 11088
SEQ ID NO 1097
LENGTH: 2321
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: Clone ID: ARATH-23APR03-CLUSTER6232_1
US-10-739-930-1097

Alignment Scores:
Pred. No.: 1.21e+03 Length: 2321
Score: 35.90 Matches: 7
Percent Similarity: 26.92% Conservative: 0
Best Local Similarity: 26.92% Mismatches: 8
Query Match: 15.47% Indels: 11
DB: 20 Gaps: 1

SEQ139-143-144-146-147-16 (1-102) x US-10-739-930-1097 (1-2321)

Qy 62 AspAsp***Leu*****Trp***** 81
Db 280 GACGACTCTCTATCTGCCACGTGG-----CGG 306

Qy 82 PhePheTrp***ThrGlu 87
Db 307 CTCCTGTGGACGACGGAG 324

RESULT 6
US-10-437-963-9511/c
Sequence 9511, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 9511

C 86 32 13.8 2740 4 US-08-756-031-4 Sequence 4, Appli
C 87 32 13.8 3373 2 US-08-897-443-2 Sequence 2, Appli
C 88 32 13.8 3438 4 US-09-949-016-5648 Sequence 5648, Ap
C 89 32 13.8 3438 4 US-09-949-016-5649 Sequence 5649, Ap
C 90 32 13.8 3449 4 US-09-907-794A-33 Sequence 33, Appl
C 91 32 13.8 3449 4 US-09-905-125A-33 Sequence 33, Appl
C 92 32 13.8 3449 4 US-09-902-775A-33 Sequence 33, Appl
C 93 32 13.8 3449 4 US-09-906-700-33 Sequence 33, Appl
C 94 32 13.8 3449 4 US-09-903-603A-33 Sequence 33, Appl
C 95 32 13.8 3449 4 US-09-904-920A-33 Sequence 33, Appl
C 96 32 13.8 3449 4 US-09-909-064-33 Sequence 33, Appl
C 97 32 13.8 3449 4 US-09-905-381A-33 Sequence 33, Appl
C 98 32 13.8 3449 4 US-09-906-618-33 Sequence 33, Appl
C 99 32 13.8 3496 4 US-09-949-016-344 Sequence 344, App
C 100 32 13.8 4335 4 US-09-949-016-15632 Sequence 15632, A

ALIGNMENTS

RESULT 1
US-09-949-016-25892
; Sequence 25892, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25892
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-25892

Alignment Scores:
Pred. No.: 24.8 Length: 601
Score: 37.00 Matches: 6
Percent Similarity: 28.57% Conservative: 0
Best Local Similarity: 28.57% Mismatches: 15
Query Match: 15.95% Indels: 0
DB: Gaps: 0

SEQ139-143-144-146-147-16 (1-102) x US-09-949-016-25892 (1-601)

Qy 82 PhePheTrp***ThrGlu*****Arg***** 101
Db 313 TTTTITGGAGACAGATCTCGCTCTGTTGCCAGGCTGGAGTGCAGCGCAGATCT 372

Qy 102 Trp 102
Db 373 TGG 375

RESULT 2
US-09-949-016-61322
; Sequence 61322, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61322
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-61322

Alignment Scores:
Pred. No.: 24.8 Length: 601
Score: 37.00 Matches: 6
Percent Similarity: 28.57% Conservative: 0
Best Local Similarity: 28.57% Mismatches: 15
Query Match: 15.95% Indels: 0
DB: Gaps: 0

SEQ139-143-144-146-147-16 (1-102) x US-09-949-016-61322 (1-601)

Qy 82 PhePheTrp***ThrGlu*****Arg***** 101
Db 313 TTTTITGGAGACAGATCTCGCTCTGTTGCCAGGCTGGAGTGCAGCGCAGATCT 372

Qy 102 Trp 102
Db 373 TGG 375

RESULT 3
US-09-949-016-12144
; Sequence 12144, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12144
; LENGTH: 47199
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12144

Alignment Scores:
Pred. No.: 3.19e+03 Length: 47199
Score: 37.00 Matches: 6
Percent Similarity: 28.57% Conservative: 0
Best Local Similarity: 28.57% Mismatches: 15
Query Match: 15.95% Indels: 0
DB: Gaps: 0

SEQ139-143-144-146-147-16 (1-102) x US-09-949-016-12144 (1-47199)

Qy 82 PhePheTrp***ThrGlu*****Arg***** 101
Db 24430 TTTTITGGAGACAGATCTCGCTCTGTTGCCAGGCTGGAGTGCAGCGCAGATCT 24489

Qy 102 Trp 102
Db 373 TGG 375

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 25, 2005, 22:14:14 ; Search time 572.195 Seconds
(without alignments)
1055.258 Million cell updates/sec

Title: SEQ139-143-144-146-147-16

Perfect score: 232

Sequence: 1 XXXRXXPKXXXXRXXXXX.....FWXTXXXXXXXRXRXW 102

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.1
Ygapop 10.0 , Ygapext 0.1
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q/cgn2_1/USPTO.spool/WERS08974584/runat_25072005_102705_5052/app_query.fasta_1.526
-DB=N Geneseq 16Dec04 -QWMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=100 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=WERS08974584 @CGN 1 1 468 @runat_25072005_102705_5052 -NCPU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSBLOK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.1 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.1 -DELOP=6 -DELEXT=7

Database : N Geneseq 16Dec04:*

1: geneseqn1980s:.*
2: geneseqn1990s:.*
3: geneseqn2000s:.*
4: geneseqn2001s:.*
5: geneseqn2001bs:.*
6: geneseqn2002as:.*
7: geneseqn2002bs:.*
8: geneseqn2003as:.*
9: geneseqn2003bs:.*
10: geneseqn2003cs:.*
11: geneseqn2003ds:.*
12: geneseqn2004as:.*
13: geneseqn2004bs:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	37	15.9	25402	4	Aak68523 Human imm
2	35.9	15.5	368	6	Abq85527 Arabidops
3	35.9	15.5	434	10	Abx61310 Arabidops
4	35.9	15.5	639	6	Abz14486 Arabidops
5	35	15.1	2755	10	Ade57777 Human gen

6	35	15.1	2755	10	ADES7781	Ade57781 Human gen
7	34.1	14.7	784	10	ADD30623	Ad30623 Plant yie
8	34.1	14.7	784	12	ADI43774	Adi43774 Plant tra
9	34.1	14.7	1111	4	AAH29588	Aah29588 Drosophil
10	34	14.7	388	4	AAI91255	Aai91255 Human pol
11	34	14.7	570	10	ADP90573	Adp90573 Human hep
12	34	14.7	784	4	AAK51974	Aak51974 Human pol
13	34	14.7	827	4	AAK52958	Aak52958 Human pol
14	34	14.7	1469	3	AAAC68082	Aac68082 Human sec
15	34	14.7	3761	8	ABZ36127	Abz36127 Human sec
16	33.1	14.3	2269	4	ABL09273	Abi09273 Drosophil
17	33.1	14.3	2761	4	ABL09154	Abi09154 Drosophil
18	33.1	14.3	4803	4	ABL09272	Abi09272 Drosophil
19	33	14.2	102	10	ABX06554	Abx06554 S. pneumo
20	33	14.2	200	12	ADH00813	Adh00813 Kidney di
21	33	14.2	326	13	ADRG6386	Adrg6386 Cotton cd
22	33	14.2	376	6	ABN94533	Abn94533 Gene #103
23	33	14.2	422	6	ABN23546	Abn23546 Human ORP
24	33	14.2	488	6	ABV96225	Abv96225 Human pan
25	33	14.2	495	5	ABA19985	Abal9985 Human ner
26	33	14.2	572	13	ACN57361	Acn57361 Cotton gy
27	33	14.2	586	13	ADQ53130	Adq53130 Novel can
28	33	14.2	595	13	ADRG6385	Adrg6385 Cotton cd
29	33	14.2	649	10	ADD29745	Add29745 Mouse tum
30	33	14.2	652	5	ABV52638	Abv52638 Human pro
31	33	14.2	763	4	AAI95301	Aai95301 Human neu
32	33	14.2	876	2	AAAX30865	Aax30865 Streptoco
33	33	14.2	1110	4	AAH90776	Aah90776 CFE 80 co
34	33	14.2	1113	8	ABZ42222	Abz42222 Streptoco
35	33	14.2	1134	4	AAH90885	Aah90885 2CFE 80 c
36	33	14.2	1171	13	ADS60924	Adsg60924 Bacterial
37	33	14.2	1174	6	ABQ53835	Abq53835 Oligonuc1
38	33	14.2	1174	6	ABQ53834	Abq53834 Oligonuc1
39	33	14.2	1311	13	ADQ81928	Adq81928 Human BMP
40	33	14.2	1554	8	ABT15151	Abt15151 Pathogen
41	33	14.2	1575	6	ABN93157	Abn93157 Staphyloc
42	33	14.2	1575	13	ADS04051	Adsg04051 Staphyloc
43	33	14.2	1603	6	ABQ60836	Abq60836 FLJ10512f
44	33	14.2	1624	2	AAV90773	Aav90773 Nucleotid
45	33	14.2	1650	4	ABL14045	Abi14045 Drosophil
46	33	14.2	2226	2	AAV73925	Aav73925 Human SAH
47	33	14.2	2226	10	ADF76062	Adf76062 Human DNA
48	33	14.2	2385	4	ABL15036	Abi15036 Drosophil
49	33	14.2	2418	6	ABZ13673	Abz13673 Arabidops
50	33	14.2	2422	4	AAH14230	Aah14230 Human cdn
51	33	14.2	2563	2	AAV28617	Aav28617 Nucleotid
52	33	14.2	2563	10	ADB68454	Adb68454 Human DCA
53	33	14.2	2592	6	ABQ60837	Abq60837 FLJ10512f
54	33	14.2	2655	10	ADB68447	Adb68447 Human DCA
55	33	14.2	2677	13	ACN37426	Acn37426 Tumour-as
56	33	14.2	2677	13	ADP23044	Adp23044 PRO polyp
57	33	14.2	2740	3	AAAC3697	Aac3697 Arabidops
58	33	14.2	2741	3	AAAC47379	Aac47379 Arabidops
59	33	14.2	2765	4	ABL28140	Abi28140 Drosophil
60	33	14.2	3616	2	AAV73924	Aav73924 Human SAH
61	33	14.2	3634	5	ABV23195	Abv23195 Human pro
62	33	14.2	3634	5	ABV29032	Abv29032 Human pro
63	33	14.2	3986	4	AAAF7821	Aaf7821 Murine TS
64	33	14.2	4126	2	AAV90882	Aav90882 Nucleotid
65	33	14.2	4292	2	AAV90772	Aav90772 Nucleotid
66	33	14.2	4333	5	ABV28571	Abv28571 Human pro
67	33	14.2	4333	5	ABV22743	Abv22743 Human pro
68	33	14.2	4333	5	ABV27685	Abv27685 Human pro
69	33	14.2	4333	5	ABV21858	Abv21858 Human pro
70	33	14.2	5201	12	ADP13477	Adp13477 Renal cel
71	33	14.2	5384	6	ABL32208	Abi32208 Human imm
72	33	14.2	6031	5	ABA15148	Abal5148 Human ner
73	33	14.2	6744	4	AAK77797	Aak77797 Human imm
74	33	14.2	6915	6	ABL32406	Abi32406 Human imm
75	33	14.2	7071	13	ADR42483	Adr42483 Aspergill
76	33	14.2	18627	2	AAV52246	Aav52246 Streptoco
77	33	14.2	44096	10	ADB68445	Adb68445 Human DCA
78	33	14.2	88853	11	ACN43904	Acn43904 Mouse gen

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 25, 2005, 22:14:49 ; Search time 3353.56 Seconds
(without alignments)
1473.786 Million cell updates/sec

Title: SEQ139-143-144-146-147-16

Perfect score: 232

Sequence: 1 XXXXXPKXXRXRXIXXXX.....FWXTXXXXXXXRXRXXXW 102

Scoring table:

	BLOSUM62	Xgapop 10.0	Xgapext 0.1
Xgapop 10.0			
Xgapext 0.1			
Fgapop 6.0			
Delop 6.0			

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/MyERS08974584/runat_25072005_102706_5062/app_query.fasta.1.526
-DB=GenEmbl -QWMT=fastcap -SUFFIX=rgc -MINWATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR SCORE=spct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=MYERS08974584 @CGN.1.1.2647 @runat_25072005_102706_5062 -NCPU=6 -ICPU=3
-NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.1 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.1 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*

2: gb.hcg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sv.*

13: gb.un.*

14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Length	ID	Description
C	1	41.9	18.1	336	8	AY201359 Arabidops
C	2	39.9	17.2	185084	5	BX629350 Zebratfish
C	3	39.9	17.2	202699	2	BX640481 Danio rer
C	4	39	16.8	104128	10	AL929562 Mouse DNA

AC130643	Rattus no	16.4	245146	2	AC130643
AC115167	Rattus no	16.4	253634	2	AC115167
AC138061	Homo sapi	15.9	106208	2	AC138061
AL139131	Homo sapi	15.9	145882	2	AL139131
AC19260	Homo sapi	15.9	153557	2	AC19260
AC129188	Mus muscu	15.9	175245	10	AC129188
AC150062	Gallus ga	15.9	175542	2	AC150062
AL390718	Homo sapi	15.9	176123	9	AL390718
AC140307	Mus muscu	15.9	187731	10	AC140307
AC150073	Gallus ga	15.9	187771	2	AC150073
AC134732	Rattus no	15.9	264975	2	AC134732
AC017708	Drosophil	15.9	37841	2	AC017708
AC006170	Drosophil	15.9	168471	3	AC006170
AE003767	Drosophil	15.9	231562	3	AE003767
BC060679	Mus muscu	15.6	4320	10	BC060679
Continuation (6 of		15.5	110000	2	AC096393
AC104098	Mus muscu	15.5	181561	2	AC104098
AC117585	Mus muscu	15.5	215745	10	AC117585
AC130088	Rattus no	15.5	231612	2	AC130088
AC131471	Rattus no	15.5	245044	2	AC131471
AC137162	Rattus no	15.5	271485	2	AC137162
AC123455	Rattus no	15.5	281085	2	AC123455
AY201051	Arabidops	15.5	573	8	AY201051
BT015452	Arabidops	15.5	591	8	BT015452
AX507596	Sequence	15.5	639	6	AX507596
AY074558	Arabidops	15.5	639	8	AY074558
AY054226	Arabidops	15.5	702	8	AY054226
BT001993	Arabidops	15.5	2259	8	BT001993
AF069299	Arabidops	15.5	95643	8	AF069299
AC12961	Rattus no	15.5	174269	2	AC12961
AL161471	Arabidops	15.5	197119	8	ATCHRIV1
AC113049	Mus muscu	15.5	199709	2	AC113049
AC134007	Rattus no	15.5	206887	2	AC134007
AC106985	Rattus no	15.5	259970	2	AC106985
AF528583	Arabidops	15.1	639	8	AF528583
AC111103	Oryza sat	15.1	222348	10	AC111103
AK067539	Oryza sat	15.1	1611	8	AK067539
AL365230	Human DNA	15.1	99257	8	AL365230
AP005879	Oryza sat	15.1	127614	8	AP005879
AP005684	Oryza sat	15.1	154912	8	AP005684
AC016037	Homo sapi	15.1	161222	2	AC016037
AC109895	Rattus no	15.1	231400	2	AC109895
AY203308	Arabidops	15.0	705	8	AY203308
AC111863	Rattus no	15.0	218729	2	AC111863
Continuation (5 of		14.9	110000	2	AC151898
Continuation (6 of		14.9	110000	2	AC151898
AC142450	Mus muscu	14.9	153108	10	AC142450
AC113662	Rattus no	14.8	237624	2	AC113662
AC095906	Rattus no	14.8	247917	2	AC095906
AF528584	Arabidops	14.7	639	8	AF528584
AF528586	Arabidops	14.7	639	8	AF528586
AF528588	Arabidops	14.7	639	8	AF528588
AF528589	Arabidops	14.7	639	8	AF528589
AF528590	Arabidops	14.7	639	8	AF528590
AV141230	Arabidops	14.7	681	8	AV141230
AX094653	Sequence	14.7	1111	6	AX094653
AC014588	Drosophil	14.7	25749	2	AC014588
AC004512	Arabidops	14.7	88292	2	AC004512
AP006583	Globebact	14.7	159119	1	AP006583
AC10666	Drosophil	14.7	175519	3	AC10666
AC102395	Mus muscu	14.7	221697	2	AC102395
AC114102	Rattus no	14.7	238352	2	AC114102
AE003599	Drosophil	14.7	250029	3	AE003599
BV201245	squm20613	14.7	201	11	BV201245
AY023258	Oryza sat	14.7	224	8	AY023258
AY692737	Saccharom	14.7	1239	8	AY692737
BD252051	47 secret	14.7	1469	6	BD252051
BC063045	Homo sapi	14.7	1700	9	BC063045
AK073936	Oryza sat	14.7	2065	8	AK073936
Z73162	S.cerevisia	14.7	2112	8	SCYLL057C
AK112076	Oryza sat	14.7	2121	8	AK112076
AK111818	Oryza sat	14.7	2353	8	AK111818
BC050388	Homo sapi	14.7	3256	9	BC050388

78	34	14.7	10849	1	AE011567	AE011567	Leptospir
79	34	14.7	37396	8	SCCH135T	Z49793	S. cerevisia
C 80	34	14.7	46189	8	AP006676	AP006676	Lotus cor
81	34	14.7	50550	7	AF271693	AF271693	Mycobacte
82	34	14.7	100899	9	AC026436	AC026436	Homo sapi
C 83	34	14.7	100898	9	AC008411	AC008411	Homo sapi
84	34	14.7	110000	1	AE017300 ¹	Continuation (2 of	
85	34	14.7	114721	8	AP004094	AP004094	Oryza sat
86	34	14.7	132906	8	CNS08C9H	AL732642	Oryza sat
87	34	14.7	139277	8	AP005115	AP005115	Oryza sat
88	34	14.7	145890	8	CNS08C70	AL731739	Oryza sat
89	34	14.7	147369	2	AC147689	AC147689	Cercopith
90	34	14.7	147512	9	AC074020	AC074020	Homo sapi
91	34	14.7	159952	9	AC007252	AC007252	Homo sapi
C 92	34	14.7	175639	10	AC133515	AC133515	Mus muscu
93	34	14.7	177675	10	AC124534	AC124534	Mus muscu
94	34	14.7	179821	2	AC147950	AC147950	Papio anu
C 95	34	14.7	183772	9	AC021006	AC021006	Homo sapi
C 96	34	14.7	192191	2	AC150426	AC150426	Branchios
97	34	14.7	203638	2	AC135953	AC135953	Macaca mu
98	34	14.7	217330	10	AC118414	AC118414	Rattus no
99	34	14.7	223729	2	AC010660	AC010660	Rattus no
100	34	14.7	237289	2	AC095523	AC095523	Rattus no

Alignment Scores:			
Pred. No.:	0.48	Length:	336
Score:	41.90	Matches:	8
Percent Similarity:	30.77%	Conservative:	0
Best Local Similarity:	30.77%	Mismatches:	7
Query Match:	18.06%	Indels:	11
DB:	8	Gaps:	1

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2005, 04:04:20 ; Search time 85 Seconds
(without alignments)
6819.690 Million cell updates/sec

Title: US-08-974-584C-118

Perfect score: 5963

Sequence: 1 MPRAPRCRAVNSLLRSYRE.....TALEAAANPALPSDFKTLTD 1132

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5952	99.8	1132	1 TERT HUMAN	O14746 homo sapien
2	5574.5	93.5	1069	2 Q8NG46	Q8NG46 homo sapien
3	4080.5	68.4	1123	1 TERT CANFA	Q6A548 canis famil
4	4046	67.9	807	2 Q8N6C3	Q8N6C3 homo sapien
5	3567	66.5	795	2 Q8NG38	Q8NG38 homo sapien
6	3619	60.7	1128	2 Q9QX24	Q9QX24 mesocricetu
7	3520.5	59.0	1125	2 Q673L6	Q673L6 rattus norv
8	3505.5	58.8	1119	2 Q673L5	Q673L5 rattus norv
9	3496	58.6	1122	1 TERT MOUSE	Q70372 mus musculu
10	2825	47.4	524	2 Q8UBR6	Q8UBR6 homo sapien
11	2581	43.3	1346	2 Q8RD80	Q8RD80 gallus gall
12	2309.5	38.7	1191	2 Q8DE32	Q8DE32 xenopus lae
13	2207.5	37.0	523	2 Q94807	Q94807 homo sapien
14	2007	33.7	575	2 Q9JK99	Q9JK99 rattus norv
15	1408	23.6	615	2 Q673L3	Q673L3 rattus norv
16	1373	22.0	514	2 Q8R266	Q8R266 mus musculu
17	717.5	12.0	1123	2 Q8SE99	Q8SE99 arabidopsis
18	717.5	12.0	1123	2 Q8SP07	Q8SP07 arabidopsis
19	683	11.5	1259	2 Q8AU13	Q8AU13 oryza sativ
20	649.5	10.9	1261	2 Q8LKW0	Q8LKW0 oryza sativ
21	588	9.9	988	1 TERT SCHPO	O13339 schizosacch
22	469	7.9	1032	2 Q7Z1L1	Q7Z1L1 euploies cr
23	468.5	7.9	939	2 Q7SD71	Q7SD71 neurospora
24	465	7.8	1032	2 Q8MUB3	Q8MUB3 euploies cr
25	450	7.5	85	2 Q8UN56	Q8UN56 homo sapien
26	449	7.5	116	2 Q8QUS5	Q8QUS5 rattus norv
27	447	7.5	940	2 Q6C9D0	Q6C9D0 yarrowia li
28	440.5	7.4	1031	1 TERT_EUPAE	O09339 euploies ae
29	432	7.2	1032	2 Q7Z1L0	Q7Z1L0 euploies cr
30	404	6.8	823	2 Q8SQQ0	Q8SQQ0 encephalito
31	397	6.7	104	2 Q9JLM1	Q9JLM1 mus musculu

ALIGNMENTS

32	394	6.6	73	2	Q9UNR4	Q9UNR4 homo sapien
33	391.5	6.6	861	2	Q6CSS0	Q6CSS0 kluyveromyc
34	388.5	6.5	1135	2	Q818Z7	Q818Z7 sterkiella
35	388	6.5	1132	1	TERT_OXYTR	O76332 oxytricha t
36	385	6.5	1108	2	Q818Z6	Q818Z6 sterkiella
37	366	6.1	867	2	Q9P8T3	Q9P8T3 candida alb
38	362	6.1	867	2	Q9P8T2	Q9P8T2 candida alb
39	359	6.0	895	2	Q9GRC5	Q9GRC5 paramescium
40	357	6.0	1117	1	TERT_TETTH	O77448 tetrahymena
41	355.5	6.0	884	1	TERT_YEAST	Q06163 saccharomyce
42	351.5	5.9	894	2	Q6BUE6	Q6BUE6 debaryomyce
43	339	5.7	79	2	Q7YR69	Q7YR69 felis silve
44	338	5.7	79	2	Q76K45	Q76K45 canis famil
45	337	5.7	896	2	Q8MUQ8	Q8MUQ8 paramescium
46	318	5.3	853	2	Q752S9	Q752S9 ashbya goss
47	310	5.2	851	2	Q6FXX4	Q6FXX4 candida gla
48	244.5	4.1	261	2	Q67EQ2	Q67EQ2 hordeum vul
49	195	3.3	260	2	Q6W8T6	Q6W8T6 euploies ra
50	194	3.3	52	2	Q9R0B3	Q9R0B3 mus musculu
51	191	3.2	233	2	Q6SXU4	Q6SXU4 euploies mi
52	191	3.2	233	2	Q6W8T4	Q6W8T4 euploies mi
53	188	3.2	960	2	Q7R365	Q7R365 giardia lam
54	187	3.1	260	2	Q6SXU6	Q6SXU6 euploies ra
55	184.5	3.1	960	2	Q9NCP5	Q9NCP5 giardia lam
56	179	3.0	2675	2	Q9FB23	Q9FB23 streptomyce
57	175.5	2.9	3530	1	MYL5_HUMAN	P03181 epatein-bar
58	172.5	2.9	660	1	YHL1_EBV	Q777A3 human herpe
59	172.5	2.9	660	2	Q777A3	Q777A3 human herpe
60	166	2.8	1474	2	Q962F9	Q962F9 cryptospori
61	161	2.7	1070	2	Q96JG5	Q96JG5 homo sapien
62	159	2.7	771	2	Q9XHY4	Q9XHY4 oryza sativ
63	158	2.6	850	2	Q66HA1	Q66HA1 rattus norv
64	154	2.6	243	2	Q6W8T3	Q6W8T3 euploies eu
65	152.5	2.6	296	2	Q69118	Q69118 human herpe
66	151.5	2.5	669	2	Q8N4X0	Q8N4X0 homo sapien
67	150.5	2.5	653	1	SF01_MOUSE	Q64213 mus musculu
68	150	2.5	611	2	Q8K0M8	Q8K0M8 mus musculu
69	150	2.5	850	2	Q80X16	Q80X16 mus musculu
70	150	2.5	850	2	Q9J315	Q9J315 mus musculu
71	149.5	2.5	886	2	Q9VDK9	Q9VDK9 mus musculu
72	148.5	2.5	1064	2	Q6P7W6	Q6P7W6 mus musculu
73	148.5	2.5	1305	2	Q8QYF9	Q8QYF9 mus musculu
74	148	2.5	243	2	Q6SXU3	Q6SXU3 euploies eu
75	147.5	2.5	986	2	Q9DW99	Q9DW99 rat cytomog
76	147	2.5	1003	1	MBD6_HUMAN	Q96DN6 homo sapien
77	147	2.5	1003	2	Q6P0P0	Q6P0P0 homo sapien
78	147	2.5	1014	2	Q6W5R1	Q6W5R1 streptomyce
79	147	2.5	1186	2	Q7T401	Q7T401 cercopithec
80	146	2.4	442	2	Q9FCJ5	Q9FCJ5 streptomyce
81	146	2.4	539	2	Q9SR71	Q9SR71 arabidopsis
82	146	2.4	953	2	Q81VT4	Q81VT4 homo sapien
83	146	2.4	1045	2	Q8N5P7	Q8N5P7 homo sapien
84	145	2.4	700	2	Q7T400	Q7T400 cercopithec
85	144	2.4	1074	2	Q6ATI5	Q6ATI5 oryza sativ
86	144	2.4	1523	2	Q93H10	Q93H10 streptomyce
87	143	2.4	1070	2	Q9APM8	Q9APM8 myxococcus
88	142.5	2.4	892	2	Q71JB1	Q71JB1 homo sapien
89	142.5	2.4	1560	2	Q60275	Q60275 homo sapien
90	142	2.4	1461	1	IE18_PRVIF	P11675 pseudorabie
91	141.5	2.4	889	2	Q9F2N5	Q9F2N5 streptomyce
92	141.5	2.4	950	2	Q82QX8	Q82QX8 streptomyce
93	141	2.4	501	2	Q8TEE9	Q8TEE9 homo sapien
94	140.5	2.4	1892	2	Q97007	Q97007 leishmania
95	140	2.3	542	2	Q68872	Q68872 myxococcus
96	140	2.3	3493	2	Q7TMR6	Q7TMR6 mus musculu
97	140	2.3	3511	1	MYL5_MOUSE	Q9QZ24 mus musculu
98	139.5	2.3	190	2	Q6W8T5	Q6W8T5 euploies va
99	139.5	2.3	552	2	Q9RSH9	Q9RSH9 deinococcus
100	139	2.3	425	2	O52231	O52231 streptococc

```
RESULT 1
TERT_HUMAN
ID TERT_HUMAN STANDARD; PRT; 1132 AA.
AC O14746; O14783;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Telomerase reverse transcriptase (EC 2.7.7.-) (Telomerase catalytic
DE subunit) (HESR2) (telomerase-associated protein 2) (TP2).
GN Name=TERT; Synonyms=EST2, TCSI, TRT;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=97400623; PubMed=9252327; DOI=10.1126/science.277.5328.955;
RA Nakamura T.M., Morin G.B., Chapman K.B., Weinrich S.L., Andrews W.H.,
RA Lingner J., Harley C.B., Cech T.R.;
RA "Telomerase catalytic subunit homologs from fission yeast and human.";
RL Science 277:955-959(1997).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97433088; PubMed=9288757; DOI=10.1016/S0092-8674(00)80538-3;
RA Meyerson M., Counter C.M., Eaton E.N., Ellisen L.W., Steiner P.,
RA Caddle S.D., Ziaugra L., Beijersbergen R.L., Davidoff M.J., Liu Q.,
RA Bacchetti S., Haber D.A., Weinberg R.A.;
RA "hEST2, the putative human telomerase catalytic subunit gene, is up-
RA regulated in tumor cells and during immortalization.";
RL Cell 90:785-795(1997).
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99267414; PubMed=10333526; DOI=10.1016/S0378-1119(99)00108-0;
RA Wick M., Zubov D., Hagen G.;
RA "Genomic organization and promoter characterization of the gene
RA encoding the human telomerase reverse transcriptase (hTERT).";
RL Gene 232:97-106(1999).
[4]
RP SEQUENCE FROM N.A.
RA Londono-Vallejo J.A.;
RA "Sequence of a BAC carrying the entire hTERT gene.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
[5]
RP FUNCTION IN TELOMERASE ACTIVITY, TISSUE SPECIFICITY, ASSOCIATION WITH
RP TEP1, AND MUTAGENESIS OF ASP-868; ASP-869 AND ASP-712.
RX PubMed=9389643;
RA Harrington L., Zhou W., McPhail T., Oulton R., Yeung D.S., Mar V.,
RA Bass M.B., Robinson M.O.;
RA "Human telomerase contains evolutionarily conserved catalytic and
RA structural subunits.";
RL Genes Dev. 11:3109-3115(1997).
[6]
RP ASSOCIATION WITH TEP1.
RX PubMed=11029039;
RA Beattie T.L., Zhou W., Robinson M.O., Harrington L.;
RA "Polymerization defects within human telomerase are distinct from
RA telomerase RNA and TEP1 binding.";
RL Mol. Biol. Cell 11:3329-3340(2000).
CC -!- FUNCTION: Telomerase is a ribonucleoprotein enzyme essential for
CC the replication of chromosome termini in most eukaryotes. It
CC elongates telomeres. It is a reverse transcriptase that adds
CC single sequence repeats to chromosome ends by copying a template
CC sequence within the RNA component of the enzyme.
CC -!- SUBUNIT: Component of the telomerase ribonucleoprotein complex at
CC least composed of TEP1, EST1A, POT1 and a telomerase RNA template
CC component (TER). Interacts with PINX1.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DISEASE: Activation of telomerase has been implicated in cell
CC immortalization and cancer cell pathogenesis.
CC -!- SIMILARITY: Belongs to the reverse transcriptase family.
CC Telomerase subfamily.

CC -!- SIMILARITY: Contains 1 reverse transcriptase domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC
CC EMBL; AF015950; AAC51672.1; -
CC EMBL; AF018167; AAC51724.1; -
CC EMBL; AF128894; AAD30037.1; -
CC EMBL; AF128893; AAD30037.1; JOINED.
CC EMBL; AY007685; AAG23289.1; -
CC PIR; T03844; T03844.
CC Genew; HGNC:11730; TERT.
CC MIM; 187270; -
CC GO; GO:0000781; C:Chromosome, telomeric region; TAS.
CC GO; GO:0042162; F:telomeric DNA binding; TAS.
CC GO; GO:0003721; F:telomeric template RNA reverse transcriptase. . .; TAS.
CC InterPro; IPR004047; RTase.
CC InterPro; IPR003545; Telomerase_RT.
CC Pfam; PF00078; RVT; 1.
CC PRINTS; PR01365; TELOMERASERT.
CC PROSITE; PS00878; RT_POL; 1.
CC DNA-binding; Nuclear protein; Ribonucleoprotein;
CC RNA-directed DNA polymerase; Telomere; Transferrase.
CC DOMAIN 605 935
CC MUTAGEN 868 868 D->A: Loss of telomerase activity.
CC MUTAGEN 868 869 DD->AA: Loss of telomerase activity.
CC MUTAGEN 869 869 D->A: Loss of telomerase activity.
CC MUTAGEN 712 712 D->A: Loss of telomerase activity.
CC CONFLICT 516 516 D -> G (in Ref. 2).
SQ SEQUENCE 1132 AA; 126996 MW; 94E35469C4CA33A0 CRC64;

Query Match 99.8%; Score 5952; DB 1; Length 1132;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1130; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPRAPRCRAVRLSLRSHYREVLPATFVRLRGLPGQWRLVQGDPAAPRALVAQCLVCPW 60
DB 1 MPRAPRCRAVRLSLRSHYREVLPATFVRLRGLPGQWRLVQGDPAAPRALVAQCLVCPW 60
QY 61 DABPPAAPSPFVQSCIKELVAVQLCERGAQNVLAFCFALLDARGGPPPEFTTSVR 120
DB 61 DABPPAAPSPFVQSCIKELVAVQLCERGAQNVLAFCFALLDARGGPPPEFTTSVR 120
QY 121 SYLPTNTVDALRGSGAWGLLRVDDVLLHLLARCALFVLVAPSCAYQVCGPPLYQLGA 180
DB 121 SYLPTNTVDALRGSGAWGLLRVDDVLLHLLARCALFVLVAPSCAYQVCGPPLYQLGA 180
QY 181 ATQARPPPHASGPRRRRLGRCERAWNHSVREAGVPLGLPAPGARRRGGSSASRLPLPKPRR 240
DB 181 ATQARPPPHASGPRRRRLGRCERAWNHSVREAGVPLGLPAPGARRRGGSSASRLPLPKPRR 240
QY 241 GAAPERTPTVGGQSWAHPTRGSDRGFCVVSPPARPAEATSEALSGSTRHSFVS 300
DB 241 GAAPERTPTVGGQSWAHPTRGSDRGFCVVSPPARPAEATSEALSGSTRHSFVS 300
QY 301 RQHHAGPSTSRPRPMDTTPCPVYATKHFLYSSGDKQLRPSFLSSRLPSLTGARRL 360
DB 301 RQHHAGPSTSRPRPMDTTPCPVYATKHFLYSSGDKQLRPSFLSSRLPSLTGARRL 360
QY 361 VETIFLGSRPMPGTPRRRLPRLPQRYWQMRPLFLELGNHACQPCPYGVLLKTHCPLRAAVT 420
DB 361 VETIFLGSRPMPGTPRRRLPRLPQRYWQMRPLFLELGNHACQPCPYGVLLKTHCPLRAAVT 420
QY 421 PAAGVCAREXPQGSVAAPSEEDTDPRLVQLLQHSFSPWQYGFVACLRRLVPPGLWS 480
DB 421 PAAGVCAREXPQGSVAAPSEEDTDPRLVQLLQHSFSPWQYGFVACLRRLVPPGLWS 480
QY 481 RHNERRFLRNTKKFISLGKAKLSLQELTWKMSVRDCAWLRRSPGVGCVPAAEHRLREI 540
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2005, 09:48:30 ; Search time 25 Seconds
(without alignments)
4356.695 Million cell updates/sec

Title: US-08-974-584C-118

Perfect score: 5963

Sequence: 1 MPRAPRCRAVRSLLRSHYRE.....TALEAANPALPSPDKFKILD 1132

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5952	99.8	1132	2 T03844	telomerase catalyt
2	717.5	12.0	1123	2 T51517	telomerase reverse
3	588.5	9.9	989	2 T03838	telomerase catalyt
4	388	6.5	1132	2 T31107	telomerase reverse
5	357	6.0	1117	2 T14891	telomerase (SC 2.7
6	355.5	6.0	884	2 S53396	telomerase catalyt
7	175.5	2.9	3530	2 A59266	unconventional myo
8	172.5	2.9	660	1 Q0BE3	BHLF1 protein - hu
9	142.5	2.4	1560	2 T00080	hypothetical prote
10	140.5	2.4	1892	2 T18314	hypothetical prote
11	140	2.3	1460	1 EDBE1F	immediate-early pr
12	140	2.3	3511	2 A59295	unconventional myo
13	139.5	2.3	552	2 F75311	ABC transporter, A
14	138.5	2.3	924	2 S27923	gene LF3 protein -
15	135	2.3	1356	1 C45219	N-methyl-D-asparta
16	134	2.2	606	2 G75302	orotidine 5'-phosp
17	133	2.2	1106	2 J00405	hypothetical 119.5
18	132.5	2.2	946	2 J07810	inositol 1,4,5-tri
19	131.5	2.2	1184	2 G01763	atrophin-1 - human
20	131	2.2	916	2 J70396	reverse transcript
21	131	2.2	1446	1 A45344	immediate-early pr
22	130.5	2.2	1184	2 S00832	atrophin-1 - human
23	130.5	2.2	2715	2 T13049	eyelid - fruit fly
24	130	2.2	1067	2 T18196	pol protein - silk
25	129.5	2.2	383	2 S32975	gene BCRC2 protein
26	129	2.2	403	2 S22796	PRP2 protein - hu
27	129	2.2	860	2 S55543	RNA-directed DNA p
28	128.5	2.2	628	2 S01955	hypothetical prote
29	128	2.1	376	2 C75580	adenine deaminase-

30	128	2.1	1776	2 G86280	protein T5E21.13 (
31	127.5	2.1	1048	2 T31425	C-terminal domain-
32	127	2.1	260	2 T23273	proline-rich prote
33	127	2.1	505	2 S72273	actin-depolymerizi
34	127	2.1	580	2 T43481	probable mucin DKF
35	127	2.1	847	1 A53800	mixed-lineage prot
36	127	2.1	862	2 T46289	hypothetical prote
37	127	2.1	1039	2 T35878	hypothetical prote
38	126.5	2.1	603	2 H75272	probable nucleic a
39	126	2.1	330	2 E98119	transposase, uncha
40	125.5	2.1	574	2 T43556	Wiskott-Aldrich sy
41	125.5	2.1	574	2 T38819	viral proteinase -
42	125	2.1	522	2 S52216	tenascin Y precurs
43	125	2.1	1914	2 T42635	immediate-early pr
44	124	2.1	1298	1 EDBE75	N-methyl-D-asparta
45	124	2.1	1323	2 S27224	hypothetical prote
46	123.5	2.1	381	2 S16506	synapsin Ia - rat
47	123.5	2.1	704	2 A30411	68.6K capsid prote
48	123	2.1	646	1 W2BEC8	salivary proline-r
49	122.5	2.1	310	1 P1HUSD	hypothetical prote
50	122.5	2.1	635	2 F75477	hypothetical prote
51	122.5	2.1	915	2 T12526	otogelin - mouse
52	122.5	2.1	2910	2 T42214	MHC class III hist
53	122	2.0	1870	2 S37671	hydroxyproline-ric
54	121.5	2.0	620	2 S06733	guanylate cyclase
55	121.5	2.0	1102	2 JH0717	period protein PER
56	121.5	2.0	1257	2 T13957	MHC class III hist
57	121	2.0	2142	2 B35098	licheninase (EC 3.
58	120.5	2.0	335	2 T05722	probable proline-r
59	120.5	2.0	891	2 G84693	ataxin-2 - mouse
60	120.5	2.0	1285	2 T14171	probable membrane
61	120	2.0	509	2 T34871	BOLFI protein - hu
62	120	2.0	1239	1 Q0BE10	DNA-directed RNA p
63	120	2.0	1859	1 A34092	DNA-directed RNA p
64	120	2.0	1862	2 T29959	MHC class III hist
65	120	2.0	1872	2 S36152	hypothetical prote
66	120	2.0	2793	2 B90784	hypothetical prote
67	120	2.0	2806	2 D85644	hypothetical prote
68	119.5	2.0	361	2 T12543	immediate-early pr
69	119.5	2.0	775	1 EDBE11	transcription fact
70	119	2.0	639	2 G02919	FK506 polyketide s
71	119	2.0	7576	2 T17428	hypothetical prote
72	118.5	2.0	628	2 S19150	hypothetical prote
73	118.5	2.0	751	2 D98320	N-Chimerin homolog
74	118.5	2.0	903	2 T00705	probable multi-dom
75	118.5	2.0	1334	2 T50568	fatty-acid synthas
76	118.5	2.0	2796	2 J04743	UL36 protein - hum
77	118.5	2.0	3164	1 WMBE66	CWI7R protein - mo
78	118	2.0	548	2 S52735	probable non-ribos
79	118	2.0	2352	2 C83229	hypothetical prote
80	117.5	2.0	312	2 A61183	probable serine/th
81	117.5	2.0	576	2 T36729	hypothetical prote
82	117.5	2.0	1222	2 T22490	collagen alpha 1(V
83	117.5	2.0	2944	2 A54849	proline-rich prote
84	117	2.0	300	2 S19560	F02569.2 protein I
85	117	2.0	512	2 E59437	probable integral
86	117	2.0	684	2 T36771	synapsin I - rat
87	117	2.0	691	2 A25704	latrophilin-1, bra
88	117	2.0	1467	2 T18411	latrophilin-1, bra
89	117	2.0	1472	2 T18413	UL43 protein - hum
90	116.5	2.0	434	1 WMBE63	probable oxidoredu
91	116.5	2.0	1174	2 B85740	probable oxidoredu
92	116.5	2.0	1174	2 H90878	hairless protein -
93	115.5	1.9	1182	2 T48378	hypothetical prote
94	115.5	1.9	1322	2 T24140	proline-rich prote
95	115	1.9	301	2 E29149	proline-rich prote
96	115	1.9	309	2 S10889	50kD proline rich
97	115	1.9	456	2 T35474	membrane transloca
98	115	1.9	629	2 E47096	probable DNA-bindi
99	115	1.9	753	2 G87178	DNA-binding protei
100	115	1.9	767	2 S41479	

protein T5E21.13 (

C-terminal domain-

proline-rich prote

actin-depolymerizi

probable mucin DKF

mixed-lineage prot

hypothetical prote

hypothetical prote

probable nucleic a

transposase, uncha

Wiskott-Aldrich sy

viral proteinase -

tenascin Y precurs

immediate-early pr

N-methyl-D-asparta

hypothetical prote

synapsin Ia - rat

68.6K capsid prote

salivary proline-r

hypothetical prote

hypothetical prote

otogelin - mouse

MHC class III hist

hydroxyproline-ric

guanylate cyclase

period protein PER

MHC class III hist

licheninase (EC 3.

probable proline-r

ataxin-2 - mouse

probable membrane

BOLFI protein - hu

DNA-directed RNA p

DNA-directed RNA p

MHC class III hist

hypothetical prote

hypothetical prote

hypothetical prote

immediate-early pr

transcription fact

FK506 polyketide s

hypothetical prote

hypothetical prote

N-Chimerin homolog

probable multi-dom

fatty-acid synthas

UL36 protein - hum

CWI7R protein - mo

probable non-ribos

hypothetical prote

probable serine/th

hypothetical prote

collagen alpha 1(V

proline-rich prote

F02569.2 protein I

probable integral

synapsin I - rat

latrophilin-1, bra

latrophilin-1, bra

UL43 protein - hum

probable oxidoredu

probable oxidoredu

hairless protein -

hypothetical prote

proline-rich prote

proline-rich prote

50kD proline rich

membrane transloca

probable DNA-bindi

DNA-binding protei

ALIGNMENTS

```
RESULT 1
telomerase catalytic chain - human
N;Alternate names: telomerase reverse transcriptase
C;Species: Homo sapiens (man)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T03844
R;Nakamura, T.M.; Morin, G.B.; Chapman, K.B.; Weinrich, S.L.; Andrews, W.H.; Lingner, J.
Science 277, 955-959, 1997
A;Title: Telomerase catalytic subunit homologs from fission yeast and human.
A;Reference number: Z15111; MUID:97400623; PMID:9252327
A;Accession: T03844
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1132 <NA>
A;Cross-references: UNIPROT:O14746; EMBL:AF015950; NID:g23330016; PIDN:AAC51672.1; PID:g2
A;Experimental source: kidney
C;Genetics:
A;Gene: TRT
A;Map position: 5p

Query Match          99.8%; Score 5952; DB 2; Length 1132;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1130; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MPAPRCRAVRSLLRSHYREVLPATFVRRLLGPGQWRVLRQGDPAAPRALVAQCLVCPW 60
Db 1 MPAPRCRAVRSLLRSHYREVLPATFVRRLLGPGQWRVLRQGDPAAPRALVAQCLVCPW 60

Qy 61 DARPPAPASPROVSCLELVARVLORLCERGAKNVAFGALLDARGGPEAFTTSVR 120
Db 61 DARPPAPASPROVSCLELVARVLORLCERGAKNVAFGALLDARGGPEAFTTSVR 120

Qy 121 SYLPTNTVDALRGSGAWGLLRVGGDDVLVHLLARCALFVLVAPSCAYQVCGPPLYOLGA 180
Db 121 SYLPTNTVDALRGSGAWGLLRVGGDDVLVHLLARCALFVLVAPSCAYQVCGPPLYOLGA 180

Qy 181 ATQARPPPHAGPRRRLCERAWNHSVREAGVPLGAPAGARRRGGSASRSLPLPKRPR 240
Db 181 ATQARPPPHAGPRRRLCERAWNHSVREAGVPLGAPAGARRRGGSASRSLPLPKRPR 240

Qy 241 GAAPEPERTVQGSWAHPGTRGSDRGFCVWSPARPAEATSLGALSSTRHSPVSG 300
Db 241 GAAPEPERTVQGSWAHPGTRGSDRGFCVWSPARPAEATSLGALSSTRHSPVSG 300

Qy 301 ROHAGPPTSRPPRPMDTPCPVVAETKHFYSSGDKQLRPSGLSSLRPSLTGARRL 360
Db 301 ROHAGPPTSRPPRPMDTPCPVVAETKHFYSSGDKQLRPSGLSSLRPSLTGARRL 360

Qy 361 VETIFLGSRPWMPGTPRRLPLQRYWQWRPLFLLELGNHAQCPYGVLLKTHCPRAAVT 420
Db 361 VETIFLGSRPWMPGTPRRLPLQRYWQWRPLFLLELGNHAQCPYGVLLKTHCPRAAVT 420

Qy 421 PAAGVCAREKQGSVAAPBEEDTDPRRLVQLLRQHSWPQVYGFVRACLRRLVPPLWGS 480
Db 421 PAAGVCAREKQGSVAAPBEEDTDPRRLVQLLRQHSWPQVYGFVRACLRRLVPPLWGS 480

Qy 481 RHNERFRUNTKFTISLGKHAQLSIQELTWKMSVRDCAWLRSPGVCVPAEHRRLREI 540
Db 481 RHNERFRUNTKFTISLGKHAQLSIQELTWKMSVRDCAWLRSPGVCVPAEHRRLREI 540

Qy 541 LAKFLHMLMSVVVVELLSFFVTTTTFQKNRLFYRPSVMSKLSQIGIRQHLKRVQLRE 600
Db 541 LAKFLHMLMSVVVVELLSFFVTTTTFQKNRLFYRPSVMSKLSQIGIRQHLKRVQLRE 600

Qy 601 LSEAEVRQHRARPALLTSRLRFTPKPDGLRPIVNMDDYVVGARTFRREKRAERLTSRVKA 660
Db 601 LSEAEVRQHRARPALLTSRLRFTPKPDGLRPIVNMDDYVVGARTFRREKRAERLTSRVKA 660

Qy 661 LFSVLNYERARRPGLGASVLGLDDIHRAWRTFVLVRQAQDPPPELYFVKVDVTGAYDTI 720
```

```
Db 661 LFSVLNYERARRPGLGASVLGLDDIHRAWRTFVLVRQAQDPPPELYFVKVDVTGAYDTI 720
Qy 721 QDRLTEVIASIIKPONTYCVRRYAVVOKAAHGHVRKAFKSHVSTLTDLPYMRQFVAHL 780
Db 721 QDRLTEVIASIIKPONTYCVRRYAVVOKAAHGHVRKAFKSHVSTLTDLPYMRQFVAHL 780
Qy 781 QETSPRLDADVIEOSSSNEASGLFDVFLRFMCHHAVIRGKSYVOCQGIPOQSILSTL 840
Db 781 QETSPRLDADVIEOSSSNEASGLFDVFLRFMCHHAVIRGKSYVOCQGIPOQSILSTL 840
Qy 841 LCSLCYGDMMENKLPAGIRRDGLLRVDDFLVTPHLLTHAKTFLRTLRVGPVEYGVVNL 900
Db 841 LCSLCYGDMMENKLPAGIRRDGLLRVDDFLVTPHLLTHAKTFLRTLRVGPVEYGVVNL 900
Qy 901 RKTVNPFPVEDEALGGTAFVQMPAHGLFPWCGLLLDTRTLEVDQSDYSYARTSTRASVTF 960
Db 901 RKTVNPFPVEDEALGGTAFVQMPAHGLFPWCGLLLDTRTLEVDQSDYSYARTSTRASVTF 960
Qy 961 NRGFKAGNNRRKLPFVLRKCHSLFLDLQVNSLQTVCTNIYKILLQAVRFHACVQLQP 1020
Db 961 NRGFKAGNNRRKLPFVLRKCHSLFLDLQVNSLQTVCTNIYKILLQAVRFHACVQLQP 1020
Qy 1021 FHQQVWKNPTFFLRVISDTASLCYSILKAKNAGMSLGAKGAAGPLPSEAVQWLCHQAPLL 1080
Db 1021 FHQQVWKNPTFFLRVISDTASLCYSILKAKNAGMSLGAKGAAGPLPSEAVQWLCHQAPLL 1080
Qy 1081 KLTRHRVTVYVPLGSLRTAQTLQSRKLPCTTLTALEAAANPALPSDFKTILD 1132
Db 1081 KLTRHRVTVYVPLGSLRTAQTLQSRKLPCTTLTALEAAANPALPSDFKTILD 1132

RESULT 2
T51517
telomerase reverse transcriptase - Arabidopsis thaliana
N;Alternate names: protein F5E19_190
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C;Accession: T51517
R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew
submitted to the Protein Sequence Database, August 2000
A;Reference number: Z25394
A;Accession: T51517
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1123 <SAT>
A;Cross-references: UNIPROT:Q9SPU7; EMBL:AL391147
A;Experimental source: cultivar Columbia; BAC clone F5E19
C;Genetics:
A;Map position: 5
A;Introns: 100/3; 125/3; 147/3; 185/1; 300/3; 325/1; 369/2; 414/3; 765/3; 942/2; 1033/2
A;Note: F5E19_190

Query Match          12.0%; Score 717.5; DB 2; Length 1123;
Best Local Similarity 23.7%; Pred. No. 9e-41;
Matches 293; Conservative 199; Mismatches 512; Indels 233; Gaps 39;

Qy 1 MPAPRCRAVRSLLRSHYREVLPATFV-----RRLGPGQWR-----L 38
Db 1 MPAPRCRAVRSLLRSHYREVLPATFV-----RRLGPGQWR-----L 38

Qy 39 VQCGDPAAPRALVAQCLVCPWDARPPAPASFRQVSCLELVARVLQR-----CERGAKN 95
Db 39 VQCGDPAAPRALVAQCLVCPWDARPPAPASFRQVSCLELVARVLQR-----CERGAKN 95

Qy 61 LRSDDPPIHYRKLHLHRCFV-VLHEQTPLDPSFTSWWSQREIVREIEMWQSGCD--CQN 117
Db 61 LRSDDPPIHYRKLHLHRCFV-VLHEQTPLDPSFTSWWSQREIVREIEMWQSGCD--CQN 117

Qy 96 VLAFGFALLDARGGPPPEAFTTSVRSYLPNTVTDALRGSGAWGLLRVGGDDVLVHLLAR 155
Db 96 VLAFGFALLDARGGPPPEAFTTSVRSYLPNTVTDALRGSGAWGLLRVGGDDVLVHLLAR 155

Qy 118 VICARYDKYDQS-----SPILLELT-SSSWBFLKRVGHDVWVYLQQ 159
Db 118 VICARYDKYDQS-----SPILLELT-SSSWBFLKRVGHDVWVYLQQ 159

Qy 156 CALFVLVAPSCAYQVCGPPLYQLGAAT-----QARPPPHASGPPRLRGGERAWNHSVRE 209
Db 156 CALFVLVAPSCAYQVCGPPLYQLGAAT-----QARPPPHASGPPRLRGGERAWNHSVRE 209

Qy 160 TSIFPLPLGKKHQVSGPPLCIKHRTKRLSLSHENKRKRDNDVQPPTKRQLWSSAVDDCPKD 219
Db 160 TSIFPLPLGKKHQVSGPPLCIKHRTKRLSLSHENKRKRDNDVQPPTKRQLWSSAVDDCPKD 219
```

1	5963	100.0	1132	15	US-10-325-810-344	Sequence 344, App
2	5963	100.0	1132	16	US-10-877-022-344	Sequence 344, App
3	5963	100.0	1132	16	US-10-877-022-344	Sequence 344, App
4	5963	100.0	1132	17	US-10-877-146-344	Sequence 344, App
5	5963	100.0	1407	14	US-10-044-659-334	Sequence 334, App
6	5963	100.0	1407	14	US-10-044-532-334	Sequence 334, App
7	5963	100.0	1407	15	US-10-325-810-628	Sequence 628, App
8	5963	100.0	1407	15	US-10-877-124-628	Sequence 628, App
9	5963	100.0	1407	16	US-10-877-022-628	Sequence 628, App
10	5963	100.0	1407	17	US-10-877-146-628	Sequence 628, App
11	5952	99.8	1132	9	US-09-990-080-2	Sequence 2, Appli

85 3496 58.6 1122 16 US-10-602-441-4 Sequence 4, Appli
86 3496 58.6 1122 16 US-10-862-698-2 Sequence 2, Appli
87 3496 58.6 1122 17 US-10-794-514A-5 Sequence 5, Appli
88 3125 52.4 622 14 US-10-294-778-12 Sequence 12, Appli
89 2581 43.3 500 14 US-10-282-960-81 Sequence 81, Appli
90 2253 37.8 1131 14 US-10-295-681-52 Sequence 52, Appli
91 2178 36.5 438 14 US-10-294-778-10 Sequence 10, Appli
92 2167.5 36.3 743 16 US-10-602-441-10 Sequence 10, Appli
93 2163.5 36.3 437 14 US-10-294-778-2 Sequence 2, Appli
94 2079 34.9 564 9 US-08-843-676-101 Sequence 101, App
95 2079 34.9 564 9 US-09-766-253-101 Sequence 101, App
96 2079 34.9 564 10 US-09-438-486-101 Sequence 101, App
97 2079 34.9 564 14 US-10-053-758-101 Sequence 101, App
98 2079 34.9 564 14 US-10-054-295-101 Sequence 101, App
99 2079 34.9 564 14 US-10-054-611-101 Sequence 101, App
100 2079 34.9 564 15 US-10-325-810-267 Sequence 267, App

ALIGNMENTS

RESULT 1
US-10-325-810-344
; Sequence 344, Application US/10325810
; Publication No. US20030204069A1
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; Linger, Joachim
; Nakamura, Toru
; Chapman, Karen B.
; Morin, Gregg B.
; Harley, Calvin B.
; Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 633
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/325,810
; FILING DATE: 20-Dec-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/402,181
; FILING DATE: 29-Sep-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:

NAME: Ausenhus, Scott L.
REGISTRATION NUMBER: 42,271
REFERENCE/DOCKET NUMBER: 015389-002620US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 344:
SEQUENCE CHARACTERISTICS:
LENGTH: 1132 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 344:
US-10-325-810-344
Query Match 100.0%; Score 5963; DB 15; Length 1132;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MPRAPRCRAVRSLLRSHYREVLPPLATEFVRLPGQWRLVORGDPAAAFRALVAQCLVCVPW 60
Db 1 MPRAPRCRAVRSLLRSHYREVLPPLATEFVRLPGQWRLVORGDPAAAFRALVAQCLVCVPW 60
Qy 61 DARPPPAAPSFQVSCLELVARVQLRCLERGAKNVLAFCGALLDARGGPPPEAFTTSVR 120
Db 61 DARPPPAAPSFQVSCLELVARVQLRCLERGAKNVLAFCGALLDARGGPPPEAFTTSVR 120
Qy 121 SYLPNTVTDALRGSGAWGLLLRRVGGDVLVHLLARCALFVLVAPSCAYQVCGPPLYQLGA 180
Db 121 SYLPNTVTDALRGSGAWGLLLRRVGGDVLVHLLARCALFVLVAPSCAYQVCGPPLYQLGA 180
Qy 181 ATQARPPPHASGPRRRRLGCERAWNHSVRBAGVPLGLPAPCARRRGGSASRLPLPKPRR 240
Db 181 ATQARPPPHASGPRRRRLGCERAWNHSVRBAGVPLGLPAPCARRRGGSASRLPLPKPRR 240
Qy 241 GAAPEPERTVGGQSWAHPCGTRGSDRGFCVVSPPARPAEEATSELEGALSGTRHSHPSVG 300
Db 241 GAAPEPERTVGGQSWAHPCGTRGSDRGFCVVSPPARPAEEATSELEGALSGTRHSHPSVG 300
Qy 301 ROHHAGPPSTSRPPRMDTFCPPVYATKIFLVSDDKEQLRPSFLSSLRPSLTGARRL 360
Db 301 ROHHAGPPSTSRPPRMDTFCPPVYATKIFLVSDDKEQLRPSFLSSLRPSLTGARRL 360
Qy 361 VETIFLGSRRPMPGCTPRRLPRLPQRYWQMPRLFLLELGNHAQCPYGVLLKTHCPLRAVT 420
Db 361 VETIFLGSRRPMPGCTPRRLPRLPQRYWQMPRLFLLELGNHAQCPYGVLLKTHCPLRAVT 420
Qy 421 PAAGVCAREKPGQSVAAPEEEDTDPRRLVQLLRQHSPPWQVYGFVRACLRLVPPGLWGS 480
Db 421 PAAGVCAREKPGQSVAAPEEEDTDPRRLVQLLRQHSPPWQVYGFVRACLRLVPPGLWGS 480
Qy 481 RHNERRPLNTKKFISLGKHAQLSLQBLTWMSVRDCAWLRSPGVCVCPAAEHLREEI 540
Db 481 RHNERRPLNTKKFISLGKHAQLSLQBLTWMSVRDCAWLRSPGVCVCPAAEHLREEI 540
Qy 541 LAKFLHMLSVVYVELLSRPFYVTTETTFQKNRLFYRPSVMSKLQSIGIQLKRVQURE 600
Db 541 LAKFLHMLSVVYVELLSRPFYVTTETTFQKNRLFYRPSVMSKLQSIGIQLKRVQURE 600
Qy 601 LSEAEVRQHREARPALITSRLRFLPKDGLRPIVNMDDYVVGARTFRREKRAERLTSRVKA 660
Db 601 LSEAEVRQHREARPALITSRLRFLPKDGLRPIVNMDDYVVGARTFRREKRAERLTSRVKA 660
Qy 661 LFSVLNYERARRPGLLGASVLGLDDIHRARWTFVLRVRAQDPPPELFFVKVDVTGADTTI 720
Db 661 LFSVLNYERARRPGLLGASVLGLDDIHRARWTFVLRVRAQDPPPELFFVKVDVTGADTTI 720
Qy 721 PODRLTEVIASIIKPNQTYCVRRYAVVQKAHGHVRKAFKSHVSTLLDLOPYMRQFVAHL 780
Db 721 PODRLTEVIASIIKPNQTYCVRRYAVVQKAHGHVRKAFKSHVSTLLDLOPYMRQFVAHL 780
Qy 781 QETSPLRDAVVIEQSSSLNEASSGLFDVFLRFMCHHAVIRGKSYVQCQIPQGSIISTLL 840
Db 781 QETSPLRDAVVIEQSSSLNEASSGLFDVFLRFMCHHAVIRGKSYVQCQIPQGSIISTLL 840

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2005, 10:42:15 ; Search time 27 Seconds
(without alignments)
3129.731 Million cell updates/sec

Title: US-08-974-584C-118
Perfect score: 5963
Sequence: 1 MPRAPRCRANRSLRSHYRE.....TALEAANPALPSPDKTILD 1132

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/ECTUS_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5963	100.0	1132	3	US-08-974-549A-344
2	5963	100.0	1132	4	US-09-402-181B-344
3	5963	100.0	1132	4	US-09-721-456-344
4	5963	100.0	1407	3	US-08-974-549A-628
5	5963	100.0	1407	4	US-08-912-951-334
6	5963	100.0	1407	4	US-09-402-181B-628
7	5963	100.0	1407	4	US-09-721-456-628
8	5952	99.8	1132	3	US-08-851-843A-225
9	5952	99.8	1132	3	US-08-974-549A-2
10	5952	99.8	1132	3	US-08-854-050-225
11	5952	99.8	1132	3	US-09-430-323-225
12	5952	99.8	1132	3	US-09-128-354-2
13	5952	99.8	1132	4	US-09-675-321-2
14	5952	99.8	1132	4	US-09-052-919-2
15	5952	99.8	1132	4	US-08-912-951-2
16	5952	99.8	1132	4	US-09-402-181B-2
17	5952	99.8	1132	4	US-09-721-456-2
18	5952	99.8	1132	4	US-09-953-052-2
19	5952	99.8	1132	4	US-09-042-460-3
20	5952	99.8	1132	4	US-09-949-016-6326
21	5952	99.8	1154	3	US-08-974-549A-611
22	5952	99.8	1154	4	US-08-912-951-323
23	5952	99.8	1154	4	US-09-402-181B-611
24	5952	99.8	1154	4	US-09-721-456-611
25	5952	99.8	1189	3	US-08-974-549A-613
26	5952	99.8	1189	4	US-08-912-951-325
27	5952	99.8	1189	4	US-09-402-181B-613

28	5952	99.8	1189	4	US-09-721-456-613	Sequence 613, App
29	5952	99.8	1200	3	US-08-974-549A-612	Sequence 612, App
30	5952	99.8	1200	4	US-08-912-951-324	Sequence 612, App
31	5952	99.8	1200	4	US-09-402-181B-612	Sequence 612, App
32	5952	99.8	1200	4	US-09-721-456-612	Sequence 612, App
33	5952	99.8	1285	3	US-08-974-549A-600	Sequence 610, App
34	5952	99.8	1285	4	US-08-912-951-314	Sequence 314, App
35	5952	99.8	1285	4	US-09-402-181B-600	Sequence 600, App
36	5952	99.8	1285	4	US-09-721-456-600	Sequence 600, App
37	5798	97.2	1193	4	US-09-949-016-11712	Sequence 11712, A
38	4046	67.9	807	3	US-08-974-549A-5	Sequence 5, Appli
39	4046	67.9	807	4	US-08-912-951-5	Sequence 5, Appli
40	4046	67.9	807	4	US-09-402-181B-5	Sequence 5, Appli
41	4046	67.9	807	4	US-09-721-456-5	Sequence 5, Appli
42	3810.5	63.9	1003	3	US-08-851-843A-217	Sequence 217, App
43	3810.5	63.9	1003	3	US-08-974-549A-336	Sequence 336, App
44	3810.5	63.9	1003	3	US-08-854-050-217	Sequence 217, App
45	3810.5	63.9	1003	3	US-09-430-323-217	Sequence 217, App
46	3810.5	63.9	1003	3	US-09-402-181B-336	Sequence 336, App
47	3810.5	63.9	1003	4	US-09-721-456-336	Sequence 336, App
48	3496	58.6	1122	4	US-09-042-460-2	Sequence 2, Appli
49	3125	52.4	622	4	US-09-582-924B-12	Sequence 12, Appli
50	2178	36.5	438	4	US-09-582-924B-10	Sequence 10, Appli
51	2163.5	36.3	437	4	US-09-582-924B-2	Sequence 2, Appli
52	2079	34.9	564	3	US-08-851-843A-101	Sequence 101, App
53	2079	34.9	564	3	US-08-974-549A-267	Sequence 267, App
54	2079	34.9	564	3	US-08-854-050-101	Sequence 101, App
55	2079	34.9	564	3	US-09-430-323-101	Sequence 101, App
56	2079	34.9	564	4	US-09-402-181B-267	Sequence 267, App
57	2079	34.9	564	4	US-09-721-456-267	Sequence 267, App
58	2079	34.9	564	4	US-09-766-233-101	Sequence 101, App
59	1666.5	27.9	364	4	US-09-417-485D-40	Sequence 40, Appli
60	1565	26.2	538	3	US-08-974-549A-602	Sequence 602, App
61	1565	26.2	538	4	US-08-912-951-316	Sequence 316, App
62	1565	26.2	538	4	US-09-402-181B-602	Sequence 602, App
63	1565	26.2	538	4	US-09-721-456-602	Sequence 602, App
64	1565	25.3	514	3	US-08-974-549A-605	Sequence 605, App
65	1506	25.3	514	4	US-08-912-951-319	Sequence 319, App
66	1506	25.3	514	4	US-08-402-181B-605	Sequence 605, App
67	1506	25.3	514	4	US-09-721-456-605	Sequence 605, App
68	1454	24.4	517	3	US-08-974-549A-606	Sequence 606, App
69	1454	24.4	517	4	US-08-912-951-320	Sequence 320, App
70	1454	24.4	517	4	US-09-402-181B-606	Sequence 606, App
71	1454	24.4	517	4	US-09-721-456-606	Sequence 606, App
72	1447	24.3	530	3	US-08-974-549A-603	Sequence 603, App
73	1447	24.3	530	4	US-08-912-951-317	Sequence 317, App
74	1447	24.3	530	4	US-09-402-181B-603	Sequence 603, App
75	1447	24.3	530	4	US-08-721-456-603	Sequence 603, App
76	1438	24.1	283	4	US-09-424-226-6	Sequence 6, Appli
77	1386	23.2	515	3	US-08-974-549A-604	Sequence 604, App
78	1386	23.2	515	4	US-08-912-951-318	Sequence 318, App
79	1386	23.2	515	4	US-09-402-181B-604	Sequence 604, App
80	1386	23.2	515	4	US-09-721-456-604	Sequence 604, App
81	1243.5	20.9	364	4	US-08-417-485D-41	Sequence 41, Appli
82	1090	18.3	259	3	US-08-974-549A-10	Sequence 10, Appli
83	1090	18.3	259	4	US-08-912-951-10	Sequence 10, Appli
84	1090	18.3	259	4	US-09-402-181B-10	Sequence 10, Appli
85	1090	18.3	259	4	US-09-721-456-10	Sequence 10, Appli
86	929	15.6	222	3	US-08-851-843A-202	Sequence 202, App
87	929	15.6	222	3	US-08-974-549A-321	Sequence 321, App
88	929	15.6	222	3	US-08-854-050-202	Sequence 202, App
89	929	15.6	222	3	US-09-430-323-202	Sequence 202, App
90	929	15.6	222	4	US-09-402-181B-321	Sequence 321, App
91	929	15.6	222	4	US-09-721-456-321	Sequence 321, App
92	905.5	15.2	330	3	US-08-851-843A-203	Sequence 203, App
93	905.5	15.2	330	3	US-08-974-549A-322	Sequence 322, App
94	905.5	15.2	330	3	US-08-854-050-203	Sequence 203, App
95	905.5	15.2	330	3	US-09-430-323-203	Sequence 203, App
96	905.5	15.2	330	4	US-09-402-181B-322	Sequence 322, App
97	905.5	15.2	330	4	US-09-721-456-322	Sequence 322, App
98	728	12.2	381	4	US-08-417-485D-47	Sequence 47, Appli
99	646	10.8	129	3	US-08-851-843A-67	Sequence 67, Appli
100	646	10.8	129	3	US-08-974-549A-13	Sequence 13, Appli

ALIGNMENTS

RESULT 1
US-08-974-549A-344
; Sequence 344, Application US/08974549A
; Patent No. 6166178
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17618
; FILING DATE: 01-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph Ted
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002610US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 344:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1132 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-974-549A-344

Query Match 100.0%; Score 5963; DB 3; Length 1132;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MPRAPRCRAVRLRSHYREVLPATFVRLRGQWRLVQGDPAAPRALVAQCVCVWP 60
Db 1 MPRAPRCRAVRLRSHYREVLPATFVRLRGQWRLVQGDPAAPRALVAQCVCVWP 60
Qy 61 DARPPPAAPSFQVSCIKELVARVQLRCERGAQNVLAFGFALLDARGGPPPEAFTTSVR 120
Db 61 DARPPPAAPSFQVSCIKELVARVQLRCERGAQNVLAFGFALLDARGGPPPEAFTTSVR 120
Qy 121 SYLPTNTVDALRGSGAWGLLLRRVDDVLVHLLARCALFVLVAPSCAYQVCGPPLYQLGA 180
Db 121 SYLPTNTVDALRGSGAWGLLLRRVDDVLVHLLARCALFVLVAPSCAYQVCGPPLYQLGA 180
Qy 181 ATQARPPPHASGPRRLGRCERAWNHSVREAGVPLGLPAPGARRRGGSASRSLPLPKRPRR 240
Db 181 ATQARPPPHASGPRRLGRCERAWNHSVREAGVPLGLPAPGARRRGGSASRSLPLPKRPRR 240
Qy 241 GAAPERTPVGQGSWAHPGRTGSPDRGFCVVSPPARPAEATSEALGSTRHSHPSVG 300
Db 241 GAAPERTPVGQGSWAHPGRTGSPDRGFCVVSPPARPAEATSEALGSTRHSHPSVG 300
Qy 301 RQHHAGPPSTSRPPRPMWDTCPVPYAEKTHFLYSSGDKQELRPSFLSSLRPSLTGARRL 360
Db 301 RQHHAGPPSTSRPPRPMWDTCPVPYAEKTHFLYSSGDKQELRPSFLSSLRPSLTGARRL 360
Qy 361 VETIFLGSRPMPCTPRRLPRLPQRYWQMRPLFLELLGNHAQCPYGVLLKTHCPLRAAVT 420
Db 361 VETIFLGSRPMPCTPRRLPRLPQRYWQMRPLFLELLGNHAQCPYGVLLKTHCPLRAAVT 420
Qy 421 PAAGVCAREXPQGSVAAPSEEDTPRRLVOLLQHSHPQVYGFVACILRLVPPGLWGS 480
Db 421 PAAGVCAREXPQGSVAAPSEEDTPRRLVOLLQHSHPQVYGFVACILRLVPPGLWGS 480
Qy 481 RHNERRFLRNTKKFISLGKAKLSLOELTWKMSVRDCAWLRSPGVCVPAAEHRLREEI 540
Db 481 RHNERRFLRNTKKFISLGKAKLSLOELTWKMSVRDCAWLRSPGVCVPAAEHRLREEI 540
Qy 541 LAKPLHWMMSVYVVELLSRPFYVYTTFTFQKRLFFYRPSVMSKLQSIGIRQHLKRVQIRE 600
Db 541 LAKPLHWMMSVYVVELLSRPFYVYTTFTFQKRLFFYRPSVMSKLQSIGIRQHLKRVQIRE 600
Qy 601 LSEAEVROHREARPAALTSRLRIPKPDGLRPIVNDYVVGARTFREKEAERLTSRVKA 660
Db 601 LSEAEVROHREARPAALTSRLRIPKPDGLRPIVNDYVVGARTFREKEAERLTSRVKA 660
Qy 661 LFSVLNYERARRPGLLGASVLGLDDTHRAWRTFVLVRAODPPPELVFVKVDVTGADTI 720
Db 661 LFSVLNYERARRPGLLGASVLGLDDTHRAWRTFVLVRAODPPPELVFVKVDVTGADTI 720
Qy 721 PQDRLTEVIASIIKPNQTYCVRRYAVVQKAHGHVRKAFKSHVSTLTDLQPMYQFVAHL 780
Db 721 PQDRLTEVIASIIKPNQTYCVRRYAVVQKAHGHVRKAFKSHVSTLTDLQPMYQFVAHL 780
Qy 781 QETSPLRDAVVIQSSSLNEASSGLFDVFLRFMCHAVIRGKSVYQCQIPQSGIISLTL 840
Db 781 QETSPLRDAVVIQSSSLNEASSGLFDVFLRFMCHAVIRGKSVYQCQIPQSGIISLTL 840
Qy 841 LCSLCYGD MENKLFAGIRRDGLLLRLVDDFLLVTPLTHAKTFLRVLVRGVPYEGCVVNL 900
Db 841 LCSLCYGD MENKLFAGIRRDGLLLRLVDDFLLVTPLTHAKTFLRVLVRGVPYEGCVVNL 900
Qy 901 RKTVMNPFVEDEALGGTAFVQMPAHGLFPWCGLLDTRTLLEVQSDYSYARTSTRASVTF 960
Db 901 RKTVMNPFVEDEALGGTAFVQMPAHGLFPWCGLLDTRTLLEVQSDYSYARTSTRASVTF 960

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2005, 02:52:05 ; Search time 80 Seconds
(without alignments)
5472.659 Million cell updates/sec

Title: US-08-974-584C-118

Perfect score: 5963
Sequence: 1 MFPRAPRCRAVRSLLRSHYRE.....TALEAANPALPSPDKTILD 1132

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : A_Geneseq_16Dec04:.*
1: geneseqp1980s:.*
2: geneseqp1990s:.*
3: geneseqp2000s:.*
4: geneseqp2001s:.*
5: geneseqp2002s:.*
6: geneseqp2003as:.*
7: geneseqp2003bs:.*
8: geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5963	100.0	1132	2	AAW56113 Human tel
2	5952	99.8	1132	2	AAW46957 Human tel
3	5952	99.8	1132	2	AAW90251 Human cat
4	5952	99.8	1132	2	AAW28881 Human tel
5	5952	99.8	1132	2	AAW32090 Human tel
6	5952	99.8	1132	2	AAW43621 A human t
7	5952	99.8	1132	2	AAW26580 Human tel
8	5952	99.8	1132	4	AAW64859 Heart mus
9	5952	99.8	1132	4	AAW64329 Human pro
10	5952	99.8	1132	4	AAW99930 Human tel
11	5952	99.8	1132	4	AAW82765 Human tel
12	5952	99.8	1132	5	AAW29226 Human tel
13	5952	99.8	1132	5	AAW27235 Human tel
14	5952	99.8	1132	6	AAW42384 Human tel
15	5952	99.8	1132	6	AAW42063 Human tel
16	5952	99.8	1132	6	AAW56676 Human tel
17	5952	99.8	1132	6	AAW58045 Human tel
18	5952	99.8	1132	7	AAW21420 Human TER
19	5952	99.8	1132	7	AAW272743 Human pro
20	5952	99.8	1132	8	AAW70114 hTERT pro
21	5952	99.8	1132	8	AAW90599 Human TER
22	5952	99.8	1132	8	AAW182172 Human tel
23	5952	99.8	1132	8	AAW70482 Human tel
24	5952	99.8	1154	2	AAW61350 Human tel
25	5952	99.8	1189	2	AAW47008 Glutathio

26	5946	99.7	1285	2	AAW47000
27	5945	99.7	1132	2	AAW71376 Human tel
28	5945	99.7	1132	2	AAW00627 Human tel
29	5945	99.7	1132	2	AAW00638 Truncated
30	5945	99.7	1132	2	AAW28401 Human EST
31	5945	99.7	1132	3	AAW96566 hEST2, a
32	5945	99.7	1132	7	AAW47061 Human TER
33	5945	99.7	1132	7	AAW40482 Human tel
34	5929	99.4	1405	2	AAW56101 Enhanced
35	5918	99.2	1166	2	AAW00647 Telomeras
36	5902.5	99.0	1199	2	AAW47007 Glutathio
37	5873	98.5	1120	2	AAW00641 Telomeras
38	5864	98.3	1120	2	AAW00650 Telomeras
39	5712	95.8	1150	2	AAW47006 Glutathio
40	5546	93.0	1053	2	AAW00640 Altered C
41	5507	92.4	1093	2	AAW00649 Altered C
42	5458	91.5	1041	2	AAW00652 Altered C
43	5458	91.5	1041	2	AAW00643 Altered C
44	5002	83.9	948	2	AAW00639 N-termina
45	4998	83.8	948	2	AAW00648 Truncated
46	4926	82.6	936	2	AAW00642 Truncated
47	4917	82.5	936	2	AAW00651 Truncated
48	4894	82.1	949	2	AAW61349 Human tel
49	4506	75.6	1152	8	ADG90609 TERT cons
50	4046	67.9	807	2	AAW46997 Human tel
51	4046	67.9	807	2	AAW00637 N-termina
52	4046	67.9	807	2	AAW00646 Truncated
53	3619	60.7	1128	7	ADD21416 Golden ha
54	3619	60.7	1128	8	ADG90603 Hamster T
55	3496	58.6	1122	2	AAW28579 Murine TE
56	3496	58.6	1122	8	ADG90601 Murine TE
57	3466	58.1	1122	5	ABB06711 Mouse tel
58	3238	54.3	617	2	AAW00636 N-termina
59	3154	52.9	588	2	AAW00635 N-termina
60	3138	52.6	588	2	AAW00644 N-termina
61	3125	52.4	622	2	AAW25463 Human CRT
62	3038	50.9	591	2	AAW97384 A cataly
63	2581	43.3	500	6	AAO29840 Human tel
64	2573	43.1	499	6	ABB99678 Amino aci
65	2253	37.8	1131	7	ADD21415 Frog TERT
66	2178	36.5	436	6	ABB99680 Splice va
67	2178	36.5	438	2	AAW25462 Human CRT
68	2167.5	36.3	743	8	ADG90607 Dog TERT
69	2167	36.3	463	6	ABB99679 Splice va
70	2163.5	36.3	437	2	AAW25461 Human CRT
71	2079	34.9	564	2	AAW56109 Human tel
72	2007	33.7	575	8	ADG90605 Rat TERT
73	1863	31.2	348	2	AAW00645 Truncated
74	1541	25.8	538	2	AAW47001 Glutathio
75	1506	25.3	514	2	AAW47004 Glutathio
76	1484	24.9	291	6	AAO29774 hTERT MHC
77	1447	24.3	531	2	AAW47002 Glutathio
78	1438	24.1	283	2	AAW43128 Human tel
79	1417.5	23.8	516	2	AAW47005 Glutathio
80	1338.5	22.4	514	2	AAW47003 Glutathio
81	1293	21.7	250	8	ADG85224 Human tel
82	1090	18.3	259	2	AAW46998 Human tel
83	902	15.1	174	6	AAO29775 hTERT MHC
84	708	11.9	379	4	AAE00431 Consensus
85	667	11.2	174	6	ABB99681 Splice va
86	645	10.8	131	2	AAW97385 Amino aci
87	588	9.5	988	2	AAW56107 S. pombe
88	565	9.5	816	8	ADG70135 HIV RT/HT
89	555	9.3	803	8	ADG70131 HIV RT/HT
90	555	9.3	816	8	ADG70133 HIV RT/HT
91	555	9.0	108	6	ABB99682 Splice va
92	535	9.0	100	5	ABG71628 hTERT fra
93	534	9.0	100	5	ABG71627 hTERT fra
94	503	8.4	100	5	ABG71627 hTERT fra
95	478	8.0	576	8	ADG70112 HIV-1 RT/
96	478	8.0	592	8	ADG70120 HIV RT/HT
97	478	8.0	605	8	ADG70121 HIV RT/HT
98	461	7.7	816	8	ADG70132 HIV RT/HT

99 459 7.7 456 8 ADG70126
100 459 7.7 462 8 ADG70124

Adg70126 HIV RT/ht
Adg70124 HIV RT/ht

ALIGNMENTS

RESULT 1

AAW56113
ID AAW56113 standard; protein; 1132 AA.

XX AC AAW56113;

XX 13-AUG-1998 (first entry)

XX Human telomerase reverse transcriptase protein refined sequence.

XX Human; telomerase reverse transcriptase; hTERT; TRT; diagnosis; prognosis;
KW cell proliferation; cancer; ageing; ribonucleoprotein.

XX Homo sapiens.

XX GB2317891-A.

XX 08-APR-1998.

XX 01-OCT-1997; 97GB-00020890.

XX 01-OCT-1996; 96US-00724643.

XX 18-APR-1997; 97US-00844419.

XX 25-APR-1997; 97US-00846017.

XX 06-MAY-1997; 97US-00851843.

XX 09-MAY-1997; 97US-00854050.

XX 14-AUG-1997; 97US-00911312.

XX 14-AUG-1997; 97US-00912951.

XX 14-AUG-1997; 97US-00915503.

XX (GERO-) GERON CORP.

XX (UYTE-) UNIV TECHNOLOGY CORP.

XX Cech TR, Lingner J, Nakamura T, Chapman KB, Morin GB, Harley CB;

XX Andrews WH;

XX WPI: 1998-171633/16.

XX N-PSDB; AAW22428.

XX Pure and recombinant human Telomerase Reverse Transcriptase and its
PT variants - are useful in the diagnosis, prognosis and treatment of cell
PT proliferation conditions especially cancer and ageing.

XX Example 1; Fig 74; 387pp; English.

XX The present sequence represents human telomerase reverse transcriptase
CC (hTERT), which is a ribonucleoprotein. The present invention also
CC describes the following methods: (A) determining whether a test compound
CC is a modulator of hTERT by detecting the change in hTERT recombinant
CC protein or polynucleotide, on administration of the compound; (B)
CC preparation of recombinant telomerase by contacting a protein preparation
CC of hTERT with a telomerase RNA component; (C) detection of the hTERT RNA or
CC protein in a sample by binding a relevant probe to the sample and
CC detecting the complex formed or in the case of RNA detection, amplifying
CC the product and correlating the presence of complex or amplification
CC product with presence of hTERT in the sample; and (D) increasing the
CC proliferation of a vertebrate cell by increasing hTERT expression; and (E)
CC the use of an agent that causes an increase in cell vertebrate cell
CC proliferation to create a medicament that inhibits ageing. A protein
CC preparation of hTERT and the polynucleotide encoding hTERT can be used in
CC the manufacture of medicaments for inhibiting the effect of ageing or
CC cancer. Inhibitors of telomerase activity can be used to treat conditions
CC that are associated with high telomerase activity. A protein preparation
CC of hTERT can also be used in the new methods

XX Sequence 1132 AA;

Query Match 100.0%; Score 5963; DB 2; Length 1132;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MPRAPRCRAVRSLLRSHYREVLPLATFVRRLGPOGWRLVORGDPAAFRALVAQCVCVPM	60
Db	1	MPRAPRCRAVRSLLRSHYREVLPLATFVRRLGPOGWRLVORGDPAAFRALVAQCVCVPM	60
Qy	61	DARPPPAAPSPFVSCUKELVARVLQRLCERGAKNVLAFGFALDVGARGGPEAFTTSVR	120
Db	61	DARPPPAAPSPFVSCUKELVARVLQRLCERGAKNVLAFGFALDVGARGGPEAFTTSVR	120
Qy	121	SYLPNTVTDALRGSGAMGLLRLRVGDDVLVHLLARCALFVLVAPSCAYQVCGPPLYOLGA	180
Db	121	SYLPNTVTDALRGSGAMGLLRLRVGDDVLVHLLARCALFVLVAPSCAYQVCGPPLYOLGA	180
Qy	181	ATQARPPPHASGPRRLGCRANWHSVREAGVPLGLPAPGARRRGGSASRSLPLPKPRR	240
Db	181	ATQARPPPHASGPRRLGCRANWHSVREAGVPLGLPAPGARRRGGSASRSLPLPKPRR	240
Qy	241	GAAPERTPVGGSGMAHPGRTGRGSDRGFCVSPARPAEATSLLEGALSGTRHSHSVG	300
Db	241	GAAPERTPVGGSGMAHPGRTGRGSDRGFCVSPARPAEATSLLEGALSGTRHSHSVG	300
Qy	301	RQHHAGPSTSRPPRPMWDTPCPPVYAEETHFLYSSGDKQLRPSFLSSLRPSLTGARRL	360
Db	301	RQHHAGPSTSRPPRPMWDTPCPPVYAEETHFLYSSGDKQLRPSFLSSLRPSLTGARRL	360
Qy	361	VETIFLGSRPMPGTPRRRLPRLPQRYQWMPRLFLELGNHAQCPYGVLLKTHCPLRAVT	420
Db	361	VETIFLGSRPMPGTPRRRLPRLPQRYQWMPRLFLELGNHAQCPYGVLLKTHCPLRAVT	420
Qy	421	PAAGVCAREKPOGSAVAPEEEDTPRELVLQRLHSHSPQVYGVFVRACTRELVPGLWGS	480
Db	421	PAAGVCAREKPOGSAVAPEEEDTPRELVLQRLHSHSPQVYGVFVRACTRELVPGLWGS	480
Qy	481	RHNERFLRNTKTFISLGKHAQLSLOELTWKMSVRDCAWLRRSPGVGCVAAEHRLREEI	540
Db	481	RHNERFLRNTKTFISLGKHAQLSLOELTWKMSVRDCAWLRRSPGVGCVAAEHRLREEI	540
Qy	541	LAKFLHLMSSVYVVELLRSFFYVTTFTFQNNRPFYFPPSPVMSKLOSTGIROHLKRVQRE	600
Db	541	LAKFLHLMSSVYVVELLRSFFYVTTFTFQNNRPFYFPPSPVMSKLOSTGIROHLKRVQRE	600
Qy	601	LSEAEVQHEARPAALTSRLRFTPKDGLRPIVNMDDYVVGARTFRREKRAERLTSRVKA	660
Db	601	LSEAEVQHEARPAALTSRLRFTPKDGLRPIVNMDDYVVGARTFRREKRAERLTSRVKA	660
Qy	661	LFSVLNYERARRPGLLGASVLGLDDIHRWRTFVLVRADQPPPELVFVKVDVTGAVDTI	720
Db	661	LFSVLNYERARRPGLLGASVLGLDDIHRWRTFVLVRADQPPPELVFVKVDVTGAVDTI	720
Qy	721	PDRLTEVIASIIKPONTYCVRRYAVVQAAHGVKAFKSHVSTLTDLQPMQFVAHL	780
Db	721	PDRLTEVIASIIKPONTYCVRRYAVVQAAHGVKAFKSHVSTLTDLQPMQFVAHL	780
Qy	781	QETSPLDADVIEOSSSINEASSGLFDVFLRFMCHHVRIRGKSVQCGIPQGSILSTL	840
Db	781	QETSPLDADVIEOSSSINEASSGLFDVFLRFMCHHVRIRGKSVQCGIPQGSILSTL	840
Qy	841	LCSLCYGDMENKLPAGIRRRDGLLRLVDDFLVTPHLLTHAKTFLTLVRGVPEYGCVVNL	900
Db	841	LCSLCYGDMENKLPAGIRRRDGLLRLVDDFLVTPHLLTHAKTFLTLVRGVPEYGCVVNL	900
Qy	901	RKTWNPFVEDEALGGTAFVQMPAHGLFPMCGLLDTRTLEVSQSDYSYARTSIRASVTF	960
Db	901	RKTWNPFVEDEALGGTAFVQMPAHGLFPMCGLLDTRTLEVSQSDYSYARTSIRASVTF	960
Qy	961	NRGFKAGRNRRKLFGLVRLKCHSLFLDLQVNSLOTCTNTYKILLQAYRFHACVLQLP	1020
Db	961	NRGFKAGRNRRKLFGLVRLKCHSLFLDLQVNSLOTCTNTYKILLQAYRFHACVLQLP	1020

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 26, 2005, 11:28:46 ; Search time 5471 Seconds
(without alignments)
7875.846 Million cell updates/sec

Title: US-08-974-584C-118
Perfect score: 5963
Sequence: 1 MPRAPCRAVRSLLRSHRYE.....TALEAANPALPSPDKTILD 1132

Scoring table:
BLOSUM62 Xgapop 10.0 , Xgapext 0.5
Xgapop 10.0 , Ygapext 0.5
Ygapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US08974584/runat 25072005 103113 9073/app query fasta_1.1287
-DB=EST -OPT=fastap -SUFFIX=n2p.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=BIOSUM62 -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US08974584 @CGN 1 1 3202 @runat 25072005 103113 9073 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THRRADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hcc: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gss1: *
9: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2636	44.2	1826	9 AY407349	Homo sapi
2	2159	36.2	1584	9 AY407350	Pan trogl
3	1796	30.1	1835	9 AY407351	Mus muscu
4	1403	23.5	925	4 BM453198	AGENCOURT
5	1025	17.2	851	5 BU702370	UI-M-F10
6	872	14.6	851	4 BG917907	602820830
7	773	13.0	492	4 BM624748	K-EST0096
8	765.5	12.8	664	5 BQ258274	NISC kp11
9	679	11.4	1424	3 CR688161	Tetraodon

10	667.5	11.2	649	7	CF531069	UI-M-FY0-
11	666.5	11.2	688	7	CF531121	UI-M-FY0-
12	640	10.7	389	1	AA281296	zt08g02.r
13	610	10.2	599	2	BB618671	BB618671
14	567	9.5	409	7	CN274427	170005313
15	566	9.5	866	2	BE371943	601217728
16	554	9.3	614	2	BB651920	BB651920
17	515	8.6	753	5	BU452535	603767927
18	514	8.6	775	4	BI388013	BFL26_002
19	510.5	8.6	1872	9	CL979127	OE1FC032
20	481.5	8.0	846	7	CN505902	AGENCOURT
21	478.5	8.0	880	5	BU377259	603811228
22	462	7.7	668	6	CA380121	659344 NC
23	456	7.6	724	5	EX889962	EX889962
24	446.5	7.5	632	6	CA353864	625469 NC
25	437.5	7.3	715	2	BE396925	601290610
26	434	7.3	703	5	EX886589	EX886589
27	433	7.3	344	7	CF531258	UI-M-FY0-
28	425	7.1	409	1	AA311750	EST182469
29	424	7.1	835	5	BU111946	603127372
30	419	7.0	679	2	BE396606	601289077
31	416	7.0	649	2	BE514070	601316575
32	405	6.8	610	2	BE514188	601316376
33	395.5	6.6	696	5	BU139751	603134527
34	387	6.5	343	6	BY783093	BY783093
35	384	6.4	338	6	BY784804	BY784804
36	380	6.4	336	6	BY775178	BY775178
37	377	6.3	347	2	AW244516	BR END068
38	374	6.3	326	5	BY149368	BY149368
39	373	6.3	641	8	AZ972318	2M0246F07
40	366	6.1	619	7	CK392784	K0850A03-
41	351.5	5.9	779	2	BE268183	601125261
42	348.5	5.8	875	5	BU122597	603148441
43	348.5	5.8	930	7	CO014076	EST802411
44	311	5.2	983	7	CO028055	EST806439
45	310.5	5.2	1197	9	CL948838	Q81F8B001
46	304	5.1	1023	7	CO024489	CO024489
47	291.5	4.9	774	6	CF547484	AGENCOURT
48	291	4.9	687	6	CD310596	StrPu691
49	284.5	4.8	813	5	BU224024	603798349
50	283	4.7	646	5	EX882610	EX882610
51	283	4.7	731	5	EX315053	EX315053
52	274	4.6	774	3	CR717591	Tetraodon
53	273.5	4.6	651	5	BM363763	BM363763
54	265	4.4	570	4	BM521744	BAK70F07
55	265	4.4	1455	8	CC190951	CH261-380
56	264	4.4	554	8	AQ397020	mgxb0014C
57	262	4.4	1554	8	CC190875	CH261-38M
58	261	4.4	1230	8	CC211572	CH261-14F
59	258.5	4.3	502	2	BF251764	EST19036
60	250	4.2	1020	9	CNS05F1M	Tetraodon
61	249.5	4.2	851	9	CG147626	PUPFP37TB
62	235	3.9	724	5	EX315052	EX315052
63	228.5	3.8	534	2	AW318894	UN08A02.Y
64	228.5	3.8	568	5	EX521269	EX521269
65	221.5	3.7	393	2	BF251775	EST19037
66	221	3.7	386	5	EX088059	EX088059
67	217.5	3.6	1009	9	CNS0730U	clone BAO
68	212	3.6	696	8	BM962910	odj21b08
69	208	3.5	1412	4	BM555089	AGENCOURT
70	207.5	3.5	2193	9	CL978828	OE1FC044
71	206	3.5	954	9	AG030799	Pan trogl
72	205.5	3.4	625	8	B27802	T1704TRD TA
73	205	3.4	613	2	BF597086	su83g03.Y
74	204	3.4	813	4	EG198331	RST17589
75	204	3.4	839	7	CO814083	AGENCOURT
76	201	3.4	817	7	CO014075	EST802410
77	199	3.3	2259	9	AG280420	Mus muscu
78	196	3.3	519	2	BF802688	PM4-C1007
79	194	3.3	1146	5	BQ683700	AGENCOURT
80	193.5	3.2	2868	9	CL982351	OE1FSC047
81	193	3.2	3953	3	AK046822	Mus muscu
82	192	3.2	1329	9	CL515191	SAIL_898_

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 26, 2005, 14:12:42 ; Search time 1052 Seconds
(without alignments)
6958.011 Million cell updates/sec

Title: US-08-974-584C-118
Perfect score: 5963
Sequence: 1 MFPRACRVRSLRLSHYRE.....TALEAANPALPSPDKTILD 1132

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7277826 seqs, 3233139505 residues

Total number of hits satisfying chosen parameters: 14555652

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US08974584/runat_25072005_103115_9161/app_query.fasta_1.1287
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=n2p.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=100 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=40 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US08974584 @CGN 1.1 487 @runat_25072005_103115_9161
-NCPUI=6 -ICPU=3 -NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEVTIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/2/pubna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	5963	100.0	3453	14	US-10-205-629-1	Sequence 1, Appli
2	5963	100.0	4037	17	US-10-325-810-343	Sequence 343, App
3	5963	100.0	4037	20	US-10-877-124-343	Sequence 343, App
4	5963	100.0	4037	20	US-10-877-022-343	Sequence 343, App
5	5963	100.0	4037	21	US-10-877-146-343	Sequence 343, App
6	5963	100.0	13766	16	US-10-105-616-1	Sequence 1, Appli
7	5952	99.8	3396	9	US-09-749-7288-32	Sequence 32, Appli
8	5952	99.8	3396	20	US-10-877-124-638	Sequence 638, App
9	5952	99.8	3396	20	US-10-877-124-639	Sequence 639, App
10	5952	99.8	3396	20	US-10-877-124-640	Sequence 640, App
11	5952	99.8	3396	20	US-10-877-124-641	Sequence 641, App
12	5952	99.8	3396	20	US-10-877-124-642	Sequence 642, App
13	5952	99.8	3396	20	US-10-877-022-638	Sequence 638, App
14	5952	99.8	3396	20	US-10-877-022-639	Sequence 639, App
15	5952	99.8	3396	20	US-10-877-022-640	Sequence 640, App
16	5952	99.8	3396	20	US-10-877-022-641	Sequence 641, App
17	5952	99.8	3396	20	US-10-877-022-642	Sequence 642, App
18	5952	99.8	3396	21	US-10-877-146-638	Sequence 638, App
19	5952	99.8	3396	21	US-10-877-146-639	Sequence 639, App
20	5952	99.8	3396	21	US-10-877-146-640	Sequence 640, App
21	5952	99.8	3396	21	US-10-877-146-641	Sequence 641, App
22	5952	99.8	3396	21	US-10-877-146-642	Sequence 642, App
23	5952	99.8	3399	19	US-10-384-339C-28	Sequence 28, Appli
24	5952	99.8	3451	20	US-10-877-124-721	Sequence 721, App
25	5952	99.8	3451	20	US-10-877-022-721	Sequence 721, App
26	5952	99.8	3451	21	US-10-877-146-721	Sequence 721, App
27	5952	99.8	4015	9	US-09-733-294A-3	Sequence 3, Appli
28	5952	99.8	4015	9	US-09-990-080-1	Sequence 1, Appli
29	5952	99.8	4015	9	US-09-843-676-224	Sequence 224, App
30	5952	99.8	4015	9	US-09-953-052-1	Sequence 1, Appli
31	5952	99.8	4015	14	US-10-053-758-224	Sequence 224, App
32	5952	99.8	4015	14	US-10-208-243-1	Sequence 1, Appli
33	5952	99.8	4015	14	US-10-054-295-224	Sequence 224, App
34	5952	99.8	4015	14	US-10-054-611-224	Sequence 224, App
35	5952	99.8	4015	14	US-10-105-663-1	Sequence 1, Appli
36	5952	99.8	4015	14	US-10-044-962-1	Sequence 1, Appli
37	5952	99.8	4015	15	US-10-044-539-1	Sequence 1, Appli
38	5952	99.8	4015	17	US-10-325-810-1	Sequence 1, Appli
39	5952	99.8	4015	17	US-10-388-578-1	Sequence 1, Appli
40	5952	99.8	4015	17	US-10-439-565-1	Sequence 1, Appli
41	5952	99.8	4015	19	US-10-602-441-1	Sequence 1, Appli
42	5952	99.8	4015	19	US-10-389-431-1	Sequence 1, Appli
43	5952	99.8	4015	20	US-10-877-124-1	Sequence 1, Appli
44	5952	99.8	4015	20	US-10-877-022-1	Sequence 1, Appli
45	5952	99.8	4015	21	US-10-831-266-3	Sequence 3, Appli
46	5952	99.8	4015	21	US-10-432-034-1	Sequence 1, Appli
47	5952	99.8	4015	21	US-10-831-267-3	Sequence 3, Appli
48	5952	99.8	4015	21	US-10-877-146-1	Sequence 1, Appli
49	5952	99.8	4015	21	US-10-917-853-1	Sequence 1, Appli
50	5952	99.8	4015	21	US-10-918-739-1	Sequence 1, Appli
51	5952	99.8	4015	21	US-10-637-443-1	Sequence 1, Appli
52	5952	99.8	4015	21	US-10-794-514A-4	Sequence 4, Appli
53	5952	99.8	4015	22	US-10-143-536-1	Sequence 1, Appli
54	5952	99.8	4027	21	US-10-840-455-2	Sequence 2, Appli
55	5945	99.7	4027	17	US-10-385-882-1	Sequence 1, Appli
56	5936	99.5	8742	16	US-10-105-616-6	Sequence 6, Appli
57	5769.5	96.8	3369	21	US-10-794-514A-339	Sequence 339, App
58	5584.5	93.7	3855	14	US-10-044-692-4	Sequence 4, Appli
59	5584.5	93.7	3855	15	US-10-044-539-4	Sequence 4, Appli
60	5584.5	93.7	3855	17	US-10-325-810-4	Sequence 4, Appli
61	5584.5	93.7	3855	20	US-10-877-124-4	Sequence 4, Appli
62	5584.5	93.7	3855	20	US-10-877-022-4	Sequence 4, Appli
63	5584.5	93.7	3855	21	US-10-877-146-4	Sequence 4, Appli
64	5570	93.4	4029	9	US-09-843-676-173	Sequence 173, App
65	5570	93.4	4029	10	US-09-438-486-173	Sequence 173, App

```
66 5570 93.4 4029 14 US-10-053-758-173 Sequence 173, App
67 5570 93.4 4029 14 US-10-054-295-173 Sequence 173, App
68 5570 93.4 4029 14 US-10-054-611-173 Sequence 173, App
69 5570 93.4 4029 17 US-10-325-810-292 Sequence 292, App
70 5570 93.4 4029 20 US-10-877-124-292 Sequence 292, App
71 5570 93.4 4029 21 US-10-877-124-292 Sequence 292, App
72 5570 93.4 4029 21 US-10-877-146-292 Sequence 292, App
73 5398.5 90.5 3255 21 US-10-794-514A-337 Sequence 337, App
74 3619 60.7 4170 19 US-10-602-441-5 Sequence 5, Appli
75 3496 58.6 3426 19 US-10-602-441-3 Sequence 6, Appli
76 3496 58.6 3426 21 US-10-794-514A-6 Sequence 6, Appli
77 3496 58.6 3496 20 US-10-862-698-11 Sequence 11, Appli
78 3125 52.4 1866 14 US-10-294-778-11 Sequence 43, Appli
79 2789 46.8 26414 21 US-10-840-455-43 Sequence 30, Appli
80 2789 46.8 51552 9 US-09-733-294A-30 Sequence 6, Appli
81 2782 46.7 4200 14 US-10-044-692-6 Sequence 6, Appli
82 2782 46.7 4200 15 US-10-044-539-6 Sequence 6, Appli
83 2782 46.7 4321 17 US-10-325-810-6 Sequence 6, Appli
84 2782 46.7 4335 20 US-10-877-124-6 Sequence 6, Appli
85 2782 46.7 4335 20 US-10-877-124-6 Sequence 6, Appli
86 2782 46.7 4335 21 US-10-877-146-6 Sequence 6, Appli
87 2782 46.7 15418 9 US-09-783-203-1 Sequence 1, Appli
88 2782 46.7 15418 9 US-09-994-427A-1 Sequence 1, Appli
89 2782 46.7 15418 10 US-09-995-419A-1 Sequence 1, Appli
90 2782 46.7 15418 14 US-10-141-220-1 Sequence 1, Appli
91 2782 46.7 15418 14 US-10-023-969-1 Sequence 1, Appli
92 2782 46.7 15418 14 US-10-206-447-1 Sequence 1, Appli
93 2782 46.7 15418 18 US-10-674-836-1 Sequence 1, Appli
94 2782 46.7 15418 19 US-10-811-012-1 Sequence 1, Appli
95 2661.5 44.6 2176 14 US-10-044-692-3 Sequence 3, Appli
96 2661.5 44.6 2176 15 US-10-044-539-3 Sequence 3, Appli
97 2661.5 44.6 2176 17 US-10-325-810-3 Sequence 3, Appli
98 2661.5 44.6 2176 20 US-10-877-124-3 Sequence 3, Appli
99 2661.5 44.6 2176 20 US-10-877-022-3 Sequence 3, Appli
100 2661.5 44.6 2176 21 US-10-877-146-3 Sequence 3, Appli
```

ALIGNMENTS

RESULT 1

```
US-10-205-629-1
; Sequence 1, Application US/10205629
; Publication No. US20030049236A1
; GENERAL INFORMATION:
; APPLICANT: Kassem, Moustapha
; APPLICANT: Jensen, Thomas
; APPLICANT: Rattan, Sureeh
; TITLE OF INVENTION: Immortalized Stem Cells
; FILE REFERENCE: 006148.00002
; CURRENT FILING DATE: 2002-07-26
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 60/315939
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: PA 2001 01148
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3453
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-205-629-1
```

Alignment Scores:

```
Pred. No.: 0 Length: 3453
Score: 5963.00 Matches: 1132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0
```

US-08-974-584C-118 (1-1132) x US-10-205-629-1 (1-3453)

```
Qy 1 MetProArgAlaProArgCysArgAlaValArgSerLeuLeuArgSerHisTyrArgGlu 20
Db 10 ATGCCGCGCGCTCCCGCTGCCGAGCGTGCGCTCCCTGCTGCGCAGCACCTACCGCGAG 69
Qy 21 ValLeuProLeuAlaThrPheValArgArgLeuGlyProGlnGlyTyrArgLeuValGln 40
Db 70 GTGTCGCGCTGGCGACGTTTCGTGCGCGCGCTGGGGGCCCCAGAGGCTGGCGCTGGTGAG 129
Qy 41 ArgGlyAspProAlaAlaPheArgAlaLeuValAlaGlnCysLeuValCysValProTrp 60
Db 130 CGCGGGGACCGCGGGGCTTCCTCCGCGCGCTGGTGCCGAGTGCCTGGTGTCCTCCCTGG 189
Qy 61 AspAlaArgProProProAlaAlaProSerPheArgGlnValSerCysLeuLysGluLeu 80
Db 190 GACGACGCGCGCGCGCGCGCGCGCGCTCTCTCCGCGAGGTGCTCTGCTGAAGAGGCTG 249
Qy 81 ValAlaArgValLeuGlnArgLeuCysGluArgGlyAlaLysAenValLeuAlaPheGly 100
Db 250 GTGCGCGAGTGTGTCAGAGGCTGTGCGAGCGCGCGCGAAGACGTGCTGGCTTCGGC 309
Qy 101 PheAlaLeuLeuAspGlyAlaArgGlyGlyProProGluAlaPheThrThrSerValArg 120
Db 310 TTCGCGCTGTGAGCGGGGCGCGGGGCGCGCGCGCGCGCGCTTCACACAGGCTGGCG 369
Qy 121 SerTyrLeuProAenThrValThrAspAlaLeuArgGlySerGlyAlaTyrGlyLeuLeu 140
Db 370 AGCTACCTGCGCAACACGCTGACCGACGCTGCGGGGGAGCGGGGCGTGGGGCTGCTG 429
Qy 141 LeuArgArgValGlyAspAspValLeuValHisLeuLeuAlaArgCysAlaLeuPheVal 160
Db 430 CTGCGCGCGTGGCGGACGCTGACCGACGCTGCGGGGGAGCGGGGCGTGGGGCTCTTTGT 489
Qy 161 LeuValAlaProSerCysAlaTyrGlnValCysGlyProProLeuTyrGlnLeuGlyAla 180
Db 490 CTGCTGCTCCCGCTGCGCTACCGAGTGTGCGGGCGCGCGCTGACAGCTCGGGCGCT 549
Qy 181 AlaThrGlnAlaArgProProHisAlaSerGlyProArgArgArgLeuGlyCysGlu 200
Db 550 GCCACTCAGCGCGCGCGCGCGCGCGCACGCTAGTGGACCCCGAAGGCGCTCTGGGATGCGAA 609
Qy 201 ArgAlaTrpAenHisSerValArgGluAlaGlyValProLeuGlyLeuProAlaProGly 220
Db 610 CGGGCTTGAAACCATAGCGTCAGGAGCGCGGGTCCCTCGGGCTGCCAGCCCGGGT 669
Qy 221 AlaArgArgGlyGlySerAlaSerArgSerLeuProLeuProLysArgProArgArg 240
Db 670 GCGAGGAGCGCGGGGCGGAGTGGCCAGCGAAGTCTGCCGTGCCCAAGAGGCCAGGGCT 729
Qy 241 GlyAlaAlaProGluProGluArgThrProValGlyGlnGlySerTrpAlaHisProGly 260
Db 730 GGGCTGCGCTGAGCGGAGCGGACGCGCTGGGCGAGGGGTCTTGGGCCACCGCGGC 789
Qy 261 ArgThrArgGlyProSerAspArgGlyPheCysValValSerProAlaArgProAlaGlu 280
Db 790 AGGACGCGTGGACCGAGTACCGGTGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 849
Qy 281 GluAlaThrSerLeuGlyAlaLeuSerGlyThrArgHisSerHisProSerValGly 300
Db 850 GAAGCCACCTCTTTGGAGGGTGGCGCTCTGTGGACGCGCGCACCTCCCAATCGTGGGC 909
Qy 301 ArgGlnHisAlaGlyProProSerThrSerArgProProArgProTrpAspThrPro 320
Db 910 CGCCAGCACCGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 969
Qy 321 CysProProValTyrAlaGluThrLysHisPheLeuTyrSerSerGlyAspLysGluGln 340
Db 970 TGTCCCGCGGTGTACCGCGAGACCAAGCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1029
Qy 341 LeuArgProSerPheLeuLeuSerSerLeuArgProSerLeuThrGlyAlaArgArgLeu 360
Db 1030 CTGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1089
Qy 361 ValGluThrIlePheLeuGlySerArgProTrpMetProGlyThrProArgArgLeuPro 380
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 26, 2005, 11:31:01 ; Search time 309 Seconds

(without alignments)
5994.386 Million cell updates/sec

Title: US-08-974-584C-118

Perfect score: 5963

Sequence: 1 MPRAPRCRAVRLSLRSHYRE.....TALEAAANPALSPDKTILD 1132

Scoring table: *BLOSUM62

Qa/cgn2_1/USPTO spool/US08974584/runat_25072005_103113_9083/app query.fasta_1.1287

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Qa/cgn2_1/USPTO spool/US08974584/runat_25072005_103113_9083/app query.fasta_1.1287
-DB=issued_Patents_NA -Qfmt=fastep -SUFFIX=n2p.rn1 -MINMATCH=0.1 -LOOPCLS=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=100 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US08974584 @CNG 1.1 59 @runat_25072005_103113_9083 -NCPUS=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/prodata/1/ina/5A COMB.seq:
2: /cgn2_6/prodata/1/ina/5B COMB.seq:
3: /cgn2_6/prodata/1/ina/6A COMB.seq:
4: /cgn2_6/prodata/1/ina/6B COMB.seq:
5: /cgn2_6/prodata/1/ina/PCTUS COMB.seq:
6: /cgn2_6/prodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	5963	100.0	4037	3	US-08-974-549A-343
2	5963	100.0	4037	4	US-09-402-181B-343
3	5963	100.0	4037	4	US-09-721-456-343
4	5952	99.8	3396	3	US-08-974-549A-638
5	5952	99.8	3396	3	US-08-974-549A-639
6	5952	99.8	3396	3	US-08-974-549A-640
7	5952	99.8	3396	3	US-08-974-549A-641
8	5952	99.8	3396	3	US-08-974-549A-642
9	5952	99.8	3396	4	US-09-721-456-638
10	5952	99.8	3396	4	US-09-721-456-639
11	5952	99.8	3396	4	US-09-721-456-640
12	5952	99.8	3396	4	US-09-721-456-641

13	5952	99.8	3396	4	US-09-721-456-642	Sequence 642, App
14	5952	99.8	3451	3	US-08-974-549A-721	Sequence 721, App
15	5952	99.8	3451	4	US-09-721-456-721	Sequence 721, App
16	5952	99.8	4015	3	US-08-851-843A-224	Sequence 224, App
17	5952	99.8	4015	3	US-08-974-549A-1	Sequence 1, Appli
18	5952	99.8	4015	3	US-08-854-050-224	Sequence 224, App
19	5952	99.8	4015	3	US-09-430-323-224	Sequence 224, App
20	5952	99.8	4015	3	US-09-572-423B-3	Sequence 3, Appli
21	5952	99.8	4015	3	US-09-128-354-1	Sequence 1, Appli
22	5952	99.8	4015	3	US-09-675-321-1	Sequence 1, Appli
23	5952	99.8	4015	3	US-09-052-919-1	Sequence 1, Appli
24	5952	99.8	4015	4	US-08-912-951-1	Sequence 1, Appli
25	5952	99.8	4015	4	US-09-733-294A-3	Sequence 3, Appli
26	5952	99.8	4015	4	US-09-402-181B-1	Sequence 1, Appli
27	5952	99.8	4015	4	US-09-721-456-1	Sequence 1, Appli
28	5952	99.8	4015	4	US-09-953-052-1	Sequence 1, Appli
29	5952	99.8	4015	4	US-09-465-491-1	Sequence 1, Appli
30	5952	99.8	4015	4	US-09-949-016-455	Sequence 455, App
31	5952	99.8	4015	4	US-09-601-645A-10	Sequence 10, Appl
32	5934	99.5	4016	4	US-09-949-016-5841	Sequence 5841, Ap
33	5584.5	93.7	3855	3	US-08-974-549A-4	Sequence 4, Appli
34	5584.5	93.7	3855	4	US-08-912-951-4	Sequence 4, Appli
35	5584.5	93.7	3855	4	US-09-402-181B-4	Sequence 4, Appli
36	5584.5	93.7	3855	4	US-09-721-456-4	Sequence 4, Appli
37	5570	93.4	4029	3	US-08-851-843A-173	Sequence 173, App
38	5570	93.4	4029	3	US-08-974-549A-292	Sequence 292, App
39	5570	93.4	4029	3	US-08-854-050-173	Sequence 173, App
40	5570	93.4	4029	3	US-09-430-323-173	Sequence 173, App
41	5570	93.4	4029	4	US-09-402-181B-292	Sequence 292, App
42	5570	93.4	4029	4	US-09-721-456-292	Sequence 292, App
43	3496	58.6	3496	4	US-09-042-460-1	Sequence 1, Appli
44	3125	52.4	1866	4	US-09-582-924B-11	Sequence 11, Appl
45	2789	46.8	44952	4	US-09-949-016-12197	Sequence 12197, A
46	2789	46.8	44960	4	US-09-949-016-17583	Sequence 17583, A
47	2789	46.8	51552	4	US-09-733-294A-30	Sequence 30, Appl
48	2782	46.7	4200	4	US-08-912-951-6	Sequence 6, Appli
49	2782	46.7	4321	4	US-09-402-181B-6	Sequence 6, Appli
50	2782	46.7	4335	3	US-08-974-549A-6	Sequence 6, Appli
51	2782	46.7	4335	3	US-09-721-456-6	Sequence 6, Appli
52	2782	46.7	15418	4	US-09-783-203-1	Sequence 1, Appli
53	2782	46.7	15418	4	US-09-994-427A-1	Sequence 1, Appli
54	2782	46.7	15418	4	US-09-244-438-1	Sequence 1, Appli
55	2661.5	44.6	2176	3	US-08-974-549A-3	Sequence 3, Appli
56	2661.5	44.6	2176	4	US-08-912-951-3	Sequence 3, Appli
57	2661.5	44.6	2176	4	US-09-402-181B-3	Sequence 3, Appli
58	2661.5	44.6	2176	4	US-09-721-456-3	Sequence 3, Appli
59	2582.5	43.3	2171	3	US-08-851-843A-100	Sequence 100, App
60	2582.5	43.3	2171	3	US-08-974-549A-266	Sequence 266, App
61	2582.5	43.3	2171	3	US-08-854-050-100	Sequence 100, App
62	2582.5	43.3	2171	3	US-09-430-323-100	Sequence 100, App
63	2582.5	43.3	2171	4	US-09-402-181B-266	Sequence 266, App
64	2582.5	43.3	2171	4	US-09-721-456-266	Sequence 266, App
65	2582.5	43.3	2171	4	US-09-766-253-100	Sequence 100, App
66	2178	36.5	1314	4	US-09-582-924B-9	Sequence 9, Appli
67	2163.5	36.3	1311	4	US-09-582-924B-1	Sequence 1, Appli
68	1356.5	22.7	7498	4	US-09-244-438-2	Sequence 2, Appli
69	1092	18.3	601	4	US-09-949-016-26770	Sequence 26770, A
70	808	13.6	601	4	US-09-949-016-26771	Sequence 26771, A
71	640	10.7	389	3	US-08-851-843A-62	Sequence 62, Appl
72	640	10.7	389	3	US-08-974-549A-8	Sequence 8, Appli
73	640	10.7	389	3	US-08-854-050-62	Sequence 62, Appl
74	640	10.7	389	3	US-09-430-323-62	Sequence 62, Appl
75	640	10.7	389	4	US-08-912-951-8	Sequence 8, Appli
76	640	10.7	389	4	US-09-402-181B-8	Sequence 8, Appli
77	640	10.7	389	4	US-09-721-456-8	Sequence 8, Appli
78	640	10.7	389	4	US-09-766-253-62	Sequence 62, Appl
79	453	7.6	2651	4	US-09-042-460-5	Sequence 5, Appli
80	425	7.1	409	4	US-09-733-294A-31	Sequence 31, Appl
81	418.5	7.0	3279	3	US-08-851-843A-1	Sequence 1, Appli
82	418.5	7.0	3279	3	US-08-974-549A-109	Sequence 109, App
83	418.5	7.0	3279	3	US-08-854-050-1	Sequence 1, Appli
84	418.5	7.0	3279	3	US-09-430-323-1	Sequence 1, Appli
85	418.5	7.0	3279	4	US-08-912-951-109	Sequence 109, App

86 418.5 7.0 3279 4 US-09-402-181B-109
 87 418.5 7.0 3279 4 US-09-721-456-109
 88 418.5 7.0 3279 4 US-09-766-253-1
 89 403.5 6.8 5544 3 US-08-851-843A-68
 90 403.5 6.8 5544 3 US-08-974-549A-111
 91 403.5 6.8 5544 3 US-08-854-050-68
 92 403.5 6.8 5544 3 US-09-430-323-68
 93 403.5 6.8 5544 4 US-08-912-951-111
 94 403.5 6.8 5544 4 US-09-402-181B-111
 95 403.5 6.8 5544 4 US-09-721-456-111
 96 403.5 6.8 5544 4 US-09-766-253-68
 97 360 6.0 2714 4 US-09-417-485D-1
 98 356 6.0 2714 4 US-09-417-485D-3
 99 355.5 6.0 2631 3 US-08-851-843A-66
 100 355.5 6.0 2631 3 US-08-974-549A-225

ALIGNMENTS

RESULT 1

US-08-974-549A-343
 ; Sequence 343, Application US/08974549A
 ; Patent No. 6166178
 ; GENERAL INFORMATION:
 ; APPLICANT: Cech, Thomas R.
 ; APPLICANT: Lingner, Joachim
 ; APPLICANT: Nakamura, Toru
 ; APPLICANT: Chapman, Karen B.
 ; APPLICANT: Morin, Gregg B.
 ; APPLICANT: Harley, Calvin B.
 ; APPLICANT: Andrews, William H.
 ; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
 ; NUMBER OF SEQUENCES: 727
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, Eighth Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-3834
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/974,549A
 ; FILING DATE: 19-NOV-1997
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/724,643
 ; FILING DATE: 01-OCT-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/844,419
 ; FILING DATE: 18-APR-1997
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/846,017
 ; FILING DATE: 25-APR-1997
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/851,843
 ; FILING DATE: 06-MAY-1997
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/854,050
 ; FILING DATE: 09-MAY-1997
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/911,312
 ; FILING DATE: 14-AUG-1997
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/912,951
 ; FILING DATE: 14-AUG-1997
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/915,503
 ; FILING DATE: 14-AUG-1997

; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: WO PCT/US97/17618
 ; FILING DATE: 01-OCT-1997
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: WO PCT/US97/17885
 ; FILING DATE: 01-OCT-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Apple, Randolph Ted
 ; REGISTRATION NUMBER: 36,429
 ; REFERENCE/DOCKET NUMBER: 015389-002610US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 576-0200
 ; TELEFAX: (415) 576-0300
 ; INFORMATION FOR SEQ ID NO: 343:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4037 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 56..3454
 ; OTHER INFORMATION: /note= "refined sequence of hTERT cDNA"
 ; US-08-974-549A-343
 ;
 ; Alignment Scores:
 ; Pred. No.: 0 Length: 4037
 ; Score: 5963.00 Matches: 1132
 ; Percent Similarity: 100.00% Conservative: 0
 ; Best Local Similarity: 100.00% Mismatches: 0
 ; Query Match: 100.00% Indels: 0
 ; DB: 3 Gaps: 0
 ;
 ; US-08-974-584C-118 (1-1132) x US-08-974-549A-343 (1-4037)
 ; QY 1 MetProArgAlaProArgCysArgAlaValArgSerLeuLeuArgSerHisTyrArgGlu 20
 ; DB 56 ATGCGGGCGCTCCCGCTGCCGAGCGCTGGCGTCCCTGCTGCCGAGCACTACCGCGAG 115
 ; QY 21 ValLeuProLeuAlaThrPheValArgArgLeuGlyProGlnGlyTyrArgLeuValGln 40
 ; DB 116 GTGCTGCGCTGGCGCACGTTCTGTCGCGCCCTGGCGGCCCGCCAGGCGTGGTGGTGCAG 175
 ; QY 41 ArgGlyAspProAlaAlaPheArgAlaValAlaGlnCysLeuValCysValProTrp 60
 ; DB 176 GCGGGGACCGCGCGGCTTTCCGCGCGCTGGTGGCCAGTCCCTGGTGGTGGTGGTGG 235
 ; QY 61 AspAlaArgProProAlaAlaProSerPheArgGlnValSerCysLeuLysGluLeu 80
 ; DB 236 GACGACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 295
 ; QY 81 ValAlaArgValLeuGlnArgLeuCysGluArgGlyAlaLysAenValLeuAlaPheGly 100
 ; DB 296 GTGGCCCGAGTGTCTGCAGAGGCTGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCG 355
 ; QY 101 PheAlaLeuLeuAspGlyAlaArgGlyClyProGluAlaPheThrSerValArg 120
 ; DB 356 TTCGCGCTGTGGACGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 415
 ; QY 121 SerTyrLeuProAsnThrValThrAspAlaLeuArgGlySerGlyAlaTrpGlyLeuLeu 140
 ; DB 416 AGCTACCTGCGCCCAACACGGTGACCGACATGGCGGGGAGCGGGGCGTGGGGGCTGTG 475
 ; QY 141 LeuArgArgValGlyAspValLeuValHisLeuLeuAlaArgCysAlaLeuPheVal 160
 ; DB 476 CTGCGCGCGTGGCGGACGACGCTGTGTTTTCCTGCTGGCGCGCTGCTGCTGCTG 535
 ; QY 161 LeuValAlaProSerCysAlaTyrGlnValCysGlyProProLeuTyrGlnLeuGlyAla 180
 ; DB 536 CTGGTGGCTCCGAGCTGCGCTACAGGTGGCGGGCGCGCGCGCGCGCGCGCGCGCT 595
 ; QY 181 AlaThrGlnAlaArgProProHisAlaSerGlyProArgArgLeuGlyCysGlu 200

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 26, 2005, 11:24:56 ; Search time 912 Seconds
(without alignments)
7347.749 Million cell updates/sec

Title: US-08-974-584c-118

Perfect score: 5963

Sequence: 1 MPRAPRCRAVRSLLRSYRE.....TALEAAANPALPSPDKTILD 1132

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US08974584/runat.25072005.103112.9051/app_query.fasta_1.1287
-DB=N Geneseq_16Dec04 -QFMT=fastap -SUFFIX=n2p.rng -MINMATCH=0.1 -LOOPCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=100 -DOCALLIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=pf0 -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US08974584@cgn.1.1.455 @runat.25072005.103112.9051 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_16Dec04:*

1: geneseqn1980a:*
2: geneseqn1990a:*
3: geneseqn2000a:*
4: geneseqn2001a:*
5: geneseqn2001b:*
6: geneseqn2002a:*
7: geneseqn2002b:*
8: geneseqn2003a:*
9: geneseqn2003b:*
10: geneseqn2003c:*
11: geneseqn2003d:*
12: geneseqn2004a:*
13: geneseqn2004b:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5963	100.0	3453	8	ABZ76217 Human TER
2	5963	100.0	4037	2	AAV22428 Human tel
3	5963	100.0	13766	6	AAZ46790 PGRN145 p
4	5952	99.8	3396	4	AAH48235 Heart mus
5	5952	99.8	3396	4	AAH49601 Human cat

6	5952	99.8	3396	4	AAH44366 Human tel
7	5952	99.8	3396	12	ADG70113 hTERT cod
8	5952	99.8	3399	6	ABV78144 Human tel
9	5952	99.8	3399	6	ABZ35720 Human elo
10	5952	99.8	3399	6	ABX09963 Human tel
11	5952	99.8	3399	6	ABL91685 Human pol
12	5952	99.8	3399	12	ADG90598 Human TER
13	5952	99.8	3798	2	AAV27876 Human tel
14	5952	99.8	3955	2	AAV22379 Human tel
15	5952	99.8	4015	2	AAZ00724 Human tel
16	5952	99.8	4015	2	AAZ20279 Human tel
17	5952	99.8	4015	2	AAZ30154 cDNA enco
18	5952	99.8	4015	4	AAH45901 Human hTE
19	5952	99.8	4015	6	AAZ46821 Human tel
20	5952	99.8	4015	6	ABA97534 Cancer ce
21	5952	99.8	4015	8	ACC57552 Human tel
22	5952	99.8	4015	8	ABZ22474 Human tel
23	5952	99.8	4015	8	ACC44482 Human tel
24	5952	99.8	4015	8	ABZ18391 Group III
25	5952	99.8	4015	10	ACC58039 Human tel
26	5952	99.8	4015	12	ADG85223 Human tel
27	5952	99.8	4015	12	ADI82171 Human cDN
28	5952	99.8	4015	13	ADR70481 Human tel
29	5952	99.8	4042	2	AAV72117 Human cat
30	5952	99.8	4070	6	ABL53711 Human tel
31	5947	99.7	4015	2	AAZ08150 Human tel
32	5945	99.7	3396	2	AAZ18266 Telomeras
33	5945	99.7	3964	2	AAZ18254 Human tel
34	5945	99.7	4023	2	AAV60320 Human tel
35	5945	99.7	4027	2	AAZ89424 Human EST
36	5945	99.7	4027	3	AAZ29388 hEST2, a
37	5945	99.7	4027	10	ADC47060 Human TER
38	5945	99.7	4027	10	ADE40481 Human tel
39	5936	99.5	8742	6	AAZ46793 pWGB5a pl
40	5994	98.8	3500	2	AAZ18275 Telomeras
41	5873	98.5	3918	2	AAZ18269 Telomeras
42	5873	98.5	3918	2	AAZ18278 Telomeras
43	5862	98.3	3543	8	ABZ69628 Plasmid c
44	5622	94.3	3203	2	AAZ18268 Altered C
45	5592	93.8	3323	2	AAZ18277 Altered C
46	5584.5	93.7	3855	2	AAV22382 Human tel
47	5543	93.0	3167	2	AAZ18271 Altered C
48	5543	93.0	3167	2	AAZ18280 Altered C
49	5432	91.1	7688	2	AAZ18351 Telomeras
50	5272	88.4	7797	2	AAZ18350 Telomeras
51	5002	83.9	3069	2	AAZ18267 Truncated
52	4998	83.8	3069	2	AAZ18276 Truncated
53	4926	82.6	3033	2	AAZ18270 Truncated
54	4926	82.6	3033	2	AAZ18279 Truncated
55	4901.5	82.2	7615	2	AAZ18349 Telomeras
56	4894	82.1	2848	2	AAV27872 Human tel
57	4421.5	74.1	2541	2	AAZ18265 N-termina
58	4421.5	74.1	2541	2	AAZ18274 Truncated
59	3619	60.7	3387	12	ADG90602 Hamster T
60	3496	58.6	3369	12	ADG90600 Murine TE
61	3378	56.6	3972	2	AAV72125 Human cat
62	3318	55.6	1883	2	AAZ18264 N-termina
63	3154	52.9	2031	2	AAZ18263 N-termina
64	3147	52.8	2031	2	AAZ18272 N-termina
65	3125	52.4	1866	2	AAZ88251 Human CRT
66	3038	50.9	2357	2	AAZ15923 cDNA enco
67	2789	46.8	51552	6	AAZ96607 DNA enco
68	2782	46.7	4335	2	AAV16979 Human tel
69	2782	46.7	15418	3	AAZ63785 Nucleotid
70	2782	46.7	15418	6	AAZ138601 Human TER
71	2782	46.7	15418	6	ABS54997 Lambda cl
72	2782	46.7	15418	6	AAZ138595 DNA of pl
73	2782	46.7	15418	10	ADC21253 Lambda cl
74	2718	45.6	3346	2	AAV60321 Alternati
75	2670.5	44.8	3175	2	AAZ15925 Nucleotid
76	2661.5	44.6	2176	2	AAV22380 Human tel
77	2582.5	43.3	2171	2	AAV22426 Human tel
78	2451	41.1	2089	2	AAV72126 Human cat

79 2178 36.5 1314 2 AAX88250
 80 2163.5 36.3 1311 2 AAX88243
 81 2159.5 36.2 1231 12 ADG90606
 82 2007 33.7 1728 12 ADG90604
 83 1868 33.3 1043 10 AXI8273
 84 1491 25.0 1012 2 AAV72124
 85 1484 24.9 873 9 AAL60416
 86 1456 24.4 2855 12 ADG90608
 87 1356.5 22.7 7498 3 AAX63786
 88 1356.5 22.7 7498 6 ABL50133
 89 1208.5 20.3 1096 2 AAX80994
 90 1052 17.6 949 2 AAV27875
 91 902 15.1 523 2 AAX15926
 92 898 15.1 519 9 AAL60417
 93 645 10.8 395 2 AAX15924
 94 565 9.5 2448 12 ADG70139
 95 555.5 9.3 2409 12 ADG70138
 96 555 9.3 1758 12 ADG70136
 97 478 8.0 1728 12 ADG70111
 98 461 7.7 2448 12 ADG70137
 99 453 7.6 2651 2 AAX80996
 100 443.5 7.4 734 4 AAK61529

ALIGNMENTS

RESULT 1
 ABZ76217 standard; DNA; 3453 BP.
 XX
 AC ABZ76217;
 XX
 DT 12-JUN-2003 (first entry)
 XX
 DE Human TERT coding DNA fragment.
 XX
 KW TERT; TRT; bone marrow stromal cell; MSC; telomeric repeat subunit;
 KW osteopathic; vulnary; gene therapy; human; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO2003010305-A1.
 XX
 PD 06-FEB-2003.
 XX
 PF 26-JUL-2002; 2002WO-DK000514.
 XX
 PR 27-JUL-2001; 2001DK-00001148.
 XX
 PA (ARHU-) ARHUS AMT.
 PA (UYAR-) UNIV ARHUS.
 XX
 PI Thomas GJ, Moustapha K, Suresh ISR;
 XX
 PS WPI; 2003-248078/24.
 XX
 PT Immortalizing human stem cells for treating osteoporosis by culturing
 PT human bone marrow stromal cells, transducing the cultures with a
 PT retroviral vector, and obtaining an immortal stem cell line.
 XX
 PS Disclosure; Fig 3-5; 38pp; English.
 XX
 CC The invention relates to immortalizing human stem cells. The method
 CC involves (a) culturing human bone marrow stromal cells (hMSC); (b)
 CC transducing the cultures with a retroviral vector, comprising the human
 CC telomeric repeat subunit (htrt) gene; and (c) obtaining an immortal stem
 CC cell line. The immortalized stem cell line has substantially identical
 CC characteristics and properties when compared to the bone marrow stromal
 CC cells. The method is useful for metabolic sink (lowering blood
 CC cholesterol level) treating bone loss associated with ageing, bone
 CC fracture and/or osteoporosis or for tissue engineering e.g., creating
 CC bone or cartilage matrix for treating bone or cartilage defects and non-
 CC healed fractures, creating skin matrix for treating skin defects or

CC burns, or producing and secreting growth factors, such as VEGF, PDGF or
 CC hGH. The immortalized stem cells are useful for drug delivery of
 CC compounds, such as Coagulation factor VIII, Coagulation factor IX,
 CC Erythropoietin, insulin, leptin, angiotensins/endostatsins, human growth
 CC hormone and/or interleukins, for drug testing (identifying drugs acting
 CC on mesenchymal stem cell and/or drugs enhancing osteoblasts
 CC differentiation and/or drugs enhancing chondrocytes differentiation, for
 CC gene therapy or for producing high amounts of pure protein for
 CC crystallization. The present sequence represents a human TERT coding
 CC fragment, which was incorporated into a retroviral vector
 XX
 SQ Sequence 3453 BP; 558 A; 1176 C; 1107 G; 612 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 8.33e-218 Length: 3453
 Score: 5963.00 Matches: 1132
 Percent Similarity: 100.00% Conservations: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 8 Gaps: 0
 US-08-974-584c-118 (1-1132) x ABZ76217 (1-3453)

Qy 1 MetProArgAlaProArgCysArgAlaValArgSerLeuLeuArgSerHisTyArgGlu 20
 Db 10 ATGCCGCGCGCTCCCGCTGCCGAGCGTGGCTCCCTGCTGCGCAGCCACTACCGCGAG 69
 Qy 21 ValLeuProLeuAlaThrPheValArgArgLeuGlyProGlnGlyTrpArgLeuValGln 40
 Db 70 GTGCTGCGCTGGCCACGTTGCTGCGCGCGCTGGGGCCCCAGGGCTGGCGCTGGTGCGAG 129
 Qy 41 ArgGlyAspProAlaAlaPheArgAlaLeuValAlaGlnCysLeuValCysValProTrp 60
 Db 130 CGCGGGGACCGCGCGGCTTCCGCGCGCTGGTGGCCAGTGCCTGGTGGTGGCTGGCTGG 189
 Qy 61 AspAlaArgProProProAlaAlaProSerPheArgGlnValSerCysLeuGlyGluLeu 80
 Db 190 GAGCACGCGCGCGCGCGCGCGCGCGCTTCTTCCGCGAGGTGCTCTGCTGAAGAGAGCTG 249
 Qy 81 ValAlaArgValLeuGlnArgLeuCysGluArgGlyAlaLysAsnValLeuAlaPheGly 100
 Db 250 GTGGCCGAGTGTGTCAGAGAGCTGTGCGAGCGCGCGCGAGAACGTCGTGGCTTCGGC 309
 Qy 101 PheAlaLeuLeuAspGlyAlaArgGlyGlyProGluAlaPheThrThrSerValArg 120
 Db 310 TTGCGCTGTGTCAGCGGGGCGCGCGGGGCCCCCGAGGCGCTTCACCCAGCGGTGGCG 369
 Qy 121 SerTyrLeuProAsnThrValThrAspAlaLeuArgGlySerGlyAlaTrpGlyLeuLeu 140
 Db 370 AGCTACCTGCCCAACACGTCGACGACACTGCGGGGAGCGGGGCGTGGGGCTGGCTG 429
 Qy 141 LeuArgArgValGlyAspAspValLeuValHisLeuLeuAlaArgCysAlaLeuPheVal 160
 Db 430 CTGCGCGCGTGGCGGCGAGCGTGTGCTTCCCTGCGGACGCTGCGCGCTCTTTGTG 489
 Qy 161 LeuValAlaProSerCysAlaTyrGlnValCysGlyProProLeuTyrGlnLeuGlyAla 180
 Db 490 CTGTGTGCTCCAGCTGGCGCTACAGGTGTGCGGGCGCGCGCTGTACCACTCGGCGCT 549
 Qy 181 AlaThrGlnAlaArgProProProHisAlaSerGlyProArgArgArgLeuGlyCysGlu 200
 Db 550 GCCACTCAGGCGCGCGCGCGCGCGCGCTAGTGGACCCCGAGGCGCTCTGGGATCGAA 609
 Qy 201 ArgAlaTrpAsnHisSerValArgGluAlaGlyValProLeuGlyLeuProAlaProGly 220
 Db 610 CGGGCTTGGACCACTAGCGTCAGGAGCGCGGGGTCCCTCGGGCTGCCAGCCCGGGT 669
 Qy 221 AlaArgArgGlyGlySerAlaSerArgSerLeuProLeuProLysArgProArgArg 240
 Db 670 GCGAGAGGCGCGGGGCGAGTGCAGGAGTCTGCGGTGCGGAGAGAGAGAGAGAGAGAG 729
 Qy 241 GlyAlaAlaProGluProGluArgThrProValGlyGlnGlySerTrpAlaHisProGly 260

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 26, 2005, 11:25:31 ; Search time 7769 Seconds

(without alignments)
7060.279 Million cell updates/sec

Title: US-08-974-584C-118

Perfect score: 5963

Sequence: 1 MPRAPRCRAVRSLLRSHYRE.....TALEAANPALPSPFKTILD 1132

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q/cgn2_1/USPTO spool/US08974584/runat_25072005_103112_9061/app_query.fasta_1.1287
-DB=GenEmbl -QFMT=fastcap -SUFFIX=n2p.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US08974584 @CGN 1.1.4648 @runat_25072005_103112_9061 -NCPU=6 -ICPU=3
-NO MWAP -LARGEQUERY -NEG-SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sv.*

13: gb.un.*

14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	5963	100.0	4037	6	E36819 Human telom
2	5963	100.0	4037	6	AR390496 Sequence
3	5963	100.0	4037	6	AR393110 Sequence
4	5963	100.0	4037	6	AX810378 Sequence

5	5963	100.0	4037	6	BD011070	Human tel
6	5963	100.0	13766	6	AX553919	Sequence
7	5952	99.8	3396	6	AR393358	Sequence
8	5952	99.8	3396	6	AR393359	Sequence
9	5952	99.8	3396	6	AR393360	Sequence
10	5952	99.8	3396	6	AR393361	Sequence
11	5952	99.8	3396	6	AR393362	Sequence
12	5952	99.8	3396	6	AX957656	Sequence
13	5952	99.8	3396	6	BD091553	Adult bon
14	5952	99.8	3396	6	BD094749	The cell
15	5952	99.8	3396	6	BD096291	Cells cap
16	5952	99.8	3399	6	AX481414	Sequence
17	5952	99.8	3451	6	AR393441	Sequence
18	5952	99.8	4015	6	AR104587	Sequence
19	5952	99.8	4015	6	AR175848	Sequence
20	5952	99.8	4015	6	BD218834	Method an
21	5952	99.8	4015	6	E36793	E36793 Human telom
22	5952	99.8	4015	6	AR182221	Sequence
23	5952	99.8	4015	6	AR224455	Sequence
24	5952	99.8	4015	6	AR226390	Sequence
25	5952	99.8	4015	6	AR243328	Sequence
26	5952	99.8	4015	6	AR263555	Sequence
27	5952	99.8	4015	6	AR265996	Sequence
28	5952	99.8	4015	6	AR390470	Sequence
29	5952	99.8	4015	6	AR393084	Sequence
30	5952	99.8	4015	6	AR404030	Sequence
31	5952	99.8	4015	6	AR438403	Sequence
32	5952	99.8	4015	6	AX019310	Sequence
33	5952	99.8	4015	6	AX133979	Sequence
34	5952	99.8	4015	6	AX552695	Sequence
35	5952	99.8	4015	6	AX810036	Sequence
36	5952	99.8	4015	6	BD011044	Human tel
37	5952	99.8	4015	6	BD015832	Quantitat
38	5952	99.8	4015	6	BD082985	Method fo
39	5952	99.8	4015	6	BD131727	Method fo
40	5952	99.8	4015	9	AF015950	Homo sapi
41	5952	99.8	4042	6	BD136185	Human tel
42	5952	99.8	4042	6	AX001446	Sequence
43	5952	99.8	4042	6	AX003121	Sequence
44	5952	99.8	4070	6	AX391846	Sequence
45	5950	99.8	4015	6	CQ717133	Sequence
46	5945	99.7	3396	6	BD196290	Vertebrat
47	5945	99.7	3964	6	BD196267	Vertebrat
48	5945	99.7	4027	6	BD233924	Method an
49	5945	99.7	8960	12	AF043739	Synthetic
50	5945	99.7	8960	12	AF043739	Synthetic
51	5936	99.5	8742	6	AX553924	Sequence
52	5894	98.8	3500	6	BD196299	Vertebrat
53	5894	98.8	3544	6	AX752636	Sequence
54	5873	98.5	3918	6	BD196293	Vertebrat
55	5822	97.6	4022	6	BD196302	Vertebrat
56	5585.5	93.7	3362	6	BD196292	Vertebrat
57	5584.5	93.7	3855	6	E36795	Human telom
58	5584.5	93.7	3855	6	AR243330	Sequence
59	5584.5	93.7	3855	6	AR390472	Sequence
60	5584.5	93.7	3855	6	AR393086	Sequence
61	5584.5	93.7	3855	6	AX810039	Sequence
62	5584.5	93.7	3855	6	BD011046	Human tel
63	5574.5	93.5	2210	9	AB085628	Homo sapi
64	5570	93.4	4029	6	AR104586	Sequence
65	5570	93.4	4029	6	AR175847	Sequence
66	5570	93.4	4029	6	E36818	E36818 Human telom
67	5570	93.4	4029	6	AR390495	Sequence
68	5570	93.4	4029	6	AR393109	Sequence
69	5570	93.4	4029	6	AX810327	Sequence
70	5570	93.4	4029	6	BD011069	Human tel
71	5534.5	92.8	3466	6	BD196301	Vertebrat
72	5506.5	92.3	3326	6	BD196295	Vertebrat
73	5460.5	91.6	3432	6	BD196304	Vertebrat
74	5432	91.1	7688	6	BD196307	Vertebrat
75	5272	88.4	7797	6	BD196306	Vertebrat
76	5207	87.3	3028	9	AB086379	Homo sapi
77	5128	86.0	2992	9	AB086950	Homo sapi

```

78 5005 83.9 3069 6 BD196291 Vertebrat
79 4954 83.1 3173 6 BD196300 Vertebrat
80 4926 82.6 3033 6 BD196294 Vertebrat
81 4904.5 82.2 7615 6 BD196305 Vertebrat
82 4875 81.8 3137 6 BD196303 Vertebrat
83 4421.5 74.1 2541 6 BD196289 Vertebrat
84 4370.5 73.3 2645 6 BD196298 Vertebrat
85 4301 68.4 3419 4 AF380351 Canis fam
86 3619 60.7 4170 10 AF149012 Mesocric
87 3520.5 59.0 3378 10 AY5339717 Rattus no
88 3506 58.8 2041 6 BD196288 Vertebrat
89 3505.5 58.8 3360 10 AY5339718 Rattus no
90 3496 58.6 3369 10 AF073311 Mus muscu
91 3496 58.6 3426 10 AF051911 Mus muscu
92 3496 58.6 3496 6 AR565399 Sequence
93 3455 57.9 2145 6 BD196297 Vertebrat
94 3419 57.3 3483 10 AY5339719 Rattus no
95 3378 56.6 3972 6 BD136189 Human tel
96 3378 56.6 3972 6 AX001451 Sequence
97 3360.5 56.4 3451 10 AY5339720 Rattus no
98 3233 54.2 3747 10 BC082327 Mus muscu
99 3154 52.9 2031 6 BD196287 Vertebrat
100 3125 52.4 1866 6 E27468 Novel gene

```

ALIGNMENTS

```

RESULT 1
E36819 4037 bp DNA linear PAT 18-JUN-2001
LOCUS Human telomerase catalytic subunit promoter.
DEFINITION Human telomerase catalytic subunit promoter.
ACCESSION E36819
VERSION E36819.1 GI:13022782
KEYWORDS JP 1999253177-A/27.
SOURCE unidentifed
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 4037)
AUTHORS Thomas,R.S., Jochimu,R., Toru,N., Karen,B.C., Greg,B.M.,
Calvin,B.H. and William,H.A.
TITLE Human telomerase catalytic subunit promoter
JOURNAL Patent: JP 1999253177-A 27 21-SEP-1999;
JERON CORP, UNIVERSITY TECHNOLOGY CORP
COMMENT OS Unidentified
PN JP 1999253177-A/27
PD 21-SEP-1999
PF 15-OCT-1998 JP 1998320169
PR 01-OCT-1996 US 08/724,643,18-APR-1997 US 08/844,419, PR
25-APR-1997 US 08/846,017,06-MAY-1997 US 08/851,843, PR
09-MAY-1997 US 08/854,050,14-AUG-1997 US 08/911,312, PR
14-AUG-1997 US 08/912,951,14-AUG-1997 US 08/915,503 PI THOMAS
R SECHI, JOCHIMU RINGER, TORU NAKAMURA, KAREN B CHAPMAN, PI GREG B
MORIN,
PI CALVIN B HAREI, WILLIAM H ANDREWS
PC C12N15/09, A61K31/70, A61K38/55, A61K39/395, A61K48/00,
PC C12Q1/02
PC C12Q1/48, C12Q1/68, G01N33/15, G01N33/48, G01N33/50, C07K14/47, PC
C07K16/40,
PC C12N1/19, C12N1/21, C12N5/10, C12N9/12, C12P21/08, (C12N1/19, PC
C12R1:84),
PC (C12N1/21, C12R1:19), (C12N9/12, C12R1:19), (C12N9/12, C12R1:84),
PC (C12N9/12, C12R1:91), C12N15/00, A61K37/64, C12N5/00 CC
Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS 56..3454.

```

```

FEATURES
source
1..4037
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:3264"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 3.66e-150 Length: 4037
Score: 5963.00 Matches: 1132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-08-974-584C-118 (1-1132) x E36819 (1-4037)
Qy 1 MetProArgAlaProArgCysArgAlaValArgSerLeuArgSerHisTyrArgGlu 20
Db 56 ATGCGCGCGCTCCCGCTGCGAGCGGTGCGCTCCCTGCTGCGCAGCCTACACCGCGAG 115
Qy 21 ValLeuProLeuAlaThrPheValArgArgLeuGlyProGlnGlyTTPArgLeuValGln 40
Db 116 GTGCTGCGCTGGCGCAGTTCGTGCGGCGCTGGGGGCCACAGGCTGCGGCTGGTGAG 175
Qy 41 ArgGlyAspProAlaAlaPheArgAlaLeuValAlaGlnCysLeuValCysValProTrp 60
Db 176 CGCGGGACCGCGGGCTTTCGCGCGCTGGTGGCCAGTGCCTGGTGGTGGTGGCTGG 235
Qy 61 AspAlaArgProProProAlaAlaProSerPheArgGlnValSerCysLeuGlyGluLeu 80
Db 236 GACGACGCGCGCGCGCGCTCCCTTCGCGCAGGTGCTGCTGCTGAGGAGCTG 295
Qy 81 ValAlaArgValLeuGlnArgLeuCysGluArgGlyAlaValAsnValLeuAlaPheGly 100
Db 296 GTGCGCGCGAGTGTGCGAGAGCTGTGCGAGCGCGCGCGAGAACGTCGTGGCTTCGGC 355
Qy 101 PheAlaLeuLeuAspGlyAlaArgGlyGlyProProGluAlaPheThrThrSerValArg 120
Db 356 TTCGCGCTGTGAGCGGGCGCGCGGGGCCCCCGAGGCTTCACCAACAGCGTGGCG 415
Qy 121 SerTyrLeuProAsnThrValThrAspAlaLeuArgGlySerGlyAlaTTPArgLeuLeu 140
Db 416 AGCTACCTGCGCAACAGGTGACGAGCAGCTGCGGGGAGCGGGGCTGGGGGCTGCTG 475
Qy 141 LeuArgArgValGlyAspAspValLeuValHisLeuLeuAlaArgCysAlaLeuPheVal 160
Db 476 CTGCGCGCGTGGCGAGCAGCTGCTGGTTACCTGCTGGCAGCTGCGGCTCTTTGTG 535
Qy 161 LeuValAlaProSerCysAlaTyrGlnValCysGlyProProLeuTyrGlnLeuGlyAla 180
Db 536 CTGCTGGCTCCAGCTGGGCTACAGGTGTGCGGGCGCGCGCTGTACCAAGCTCGGCG 595
Qy 181 AlaThrGlnAlaArgProProProHisAlaSerGlyProArgArgLeuGlyCysGlu 200
Db 596 GCCACTCAGGCGCGCGCGCGCGCCACAGCTAGTGGAGCCCGAGGCGCTTGGGATGCGAA 655
Qy 201 ArgAlaTrpAsnHisSerValArgGluAlaGlyValProLeuGlyLeuProAlaProGly 220
Db 656 CGGGCCCTGGAAACCATAGCTGAGGAGCGCGGGTCCCTTGGGCGCTGCCAGCCCCGGT 715
Qy 221 AlaArgArgArgGlyGlySerAlaSerArgSerLeuProLeuProLysArgProArgArg 240
Db 716 GCGAGGAGGCGGGGCGAGTGCAGCGGAAGTCTGCGCTGCCCAAGAGGCCCGGCGGT 775
Qy 241 GlyAlaAlaProGluProGluArgThrProValGlyGlnGlySerTrpAlaHisProGly 260
Db 776 GGCCTGCCCTGAGCGGAGCGGAGCGCGCGCTTGGGAGGGGTCTGTGGCCACCGCGGC 835
Qy 261 ArgThrArgGlyProSerAspArgGlyPheCysValValSerProAlaArgProAlaGlu 280
Db 836 AGGACGCTGGACGAGTACCGGTGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 895
Qy 281 GluAlaThrSerLeuGluGlyAlaLeuSerGlyThrArgHisSerHisProSerValGly 300
Db 896 GAACCACTCTTTGGAGGGTGGCTCTCTGGGACGCGGCACCTCCACCCCATCGTGGGC 955
Qy 301 ArgGlnHisAlaGlyProProSerThrSerArgProArgProArgProTrpAspThrPro 320
Db 956 CGCCAGCACGCGGGGCGCGCCCATCCACATCGCGGCGCACCCACGCTGCTGGGACACGCT 1015

```